Multiplex real-time quantitative PCR detection of Bufavirus, Tusavirus, and Cutavirus DNA

Yu Fu

Supervisor: Maria Söderlund-Venermo

Helsinki May, 2017

Master of Science

Master's Programme in Translational Medicine

UNIVERSITY OF HELSINKI
Faculty of Medicine

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Laitos — Institution — Department
Faculty of Medicine	Department of Virology

Tekijä — Författare — Author
Yu Fu

Työn nimi — Arbetets titel — Title

Multiplex real-time quantitative PCR detection of Bufavirus, Tusavirus, and Cutavirus DNA

Oppiaine — Läroämne — Subject
Translational Medicine

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Master of Science	May, 2017	48 pages

Tiivistelmä — Referat — Abstract

Parvoviruses are among the smallest known viruses with a genome of ~ 5 kilobases. To date, six parvoviruses have been identified in human samples, with parvovirus B19 and human bocavirus being the only two known human pathogens in this family. Bufavirus, tusavirus and cutavirus are the most recently discovered parvoviruses, all belonging to the *Protoparvovirus* genus. Bufavirus was predominately discovered in fecal samples of children with diarrhea in Africa, Asia, and Europe. Cutavirus was detected in fecal samples of children with diarrhea in South America and Africa but has also been found in skin biopsies of patients with cutaneous T cell carcinoma and malignant melanoma. Tusavirus was discovered in a stool specimen of one child with diarrhea in Tunisia. At the moment, there are too little data to determine the identity of tusavirus as a human virus. More data and evidence are required to assess the association of the three parvoviruses with human diseases.

In this study, a multiplex real-time PCR method was established to facilitate the detection and quantification of the three human-associated protoparvoviruses for further epidemiology and pathobiology study. Differentiation of different viruses was achieved by using three uniquely labeled probes. The multiplex assay was able to detect $\leqslant 10$ copies/μL of bufavirus, tusavirus, and cutavirus plasmid templates simultaneously, with an average efficiency from 100.54% to 103.76%. The assay was applied to assess the prevalence of the three viruses in skin tissues of 93 non-immunosuppressed individuals with contact dermatitis and 137 immunosuppressed transplant recipients. Bufavirus and tusavirus DNA was detected in neither of the cohorts, which might indicate the rarity of the two viruses in skin tissues. Cutavirus DNA was detected in four (2.92%) transplant recipients, but all samples from the non-immunosuppressed group were negative. Among the four cutavirus positive patients, one was diagnosed with squamous cell carcinoma. These findings further support previous discoveries of cutavirus DNA in skin tissues and serve as evidence for the identity of cutavirus as a human virus. However, its association with cancer remains to be further investigated.

Avainsanat — Nyckelord — Keywords
parvovirus, protoparvovirus, qPCR, multiplex, bufavirus, tusavirus, cutavirus

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — övriga uppgifter — Additional information

"The most exciting phrase to hear in science, the one that heralds the most discoveries, is not 'Eureka!' but 'That's funny'."

Isaac Asimov

Abbreviations

aa	amino acid
AAV	adeno-associated virus
ACD	allergic contact dermatitis
ARTI	acute respiratory tract infection
AVG	average
BKPyV	BK polyomavirus
BLAST	basic local alignment search tool
BuV	bufavirus
CPV	canine parvovirus
C_q	quantification cycle
CTCL	cutaneous T cell lymphomas
CuV	cutavirus
CV	coefficients of variations
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA
EDTA	ethylenediaminetetraacetic acid
FPV	feline parvovirus
FRET	fluorescence resonance energy transfer
GE	gastroenteritis
HBoV	human bocavirus
HFMD	hand-foot-and-mouth disease
HPyV	human polyomavirus

ICD	irritant contact dermatitis
ICTV	the International Committee on the Taxonomy of Viruses
IUPAC	the International Union of Pure and Applied Chemistry
JCPyV	JC polyomavirus
kb	kilobase
KIPyV	KI polyomavirus
KRV	Kilham rat virus
LOD	limit of detection
MCPyV	Merkel cell polyomavirus
MIQE	minimum information for publication of quantitative real-time PCR experiments
MWPyV	MW polyomavirus
NA	not available
ND	no data
NJPyV	New Jersey polyomavirus
nm	nanometer
NPAFP	non-polio acute flaccid paralysis
NS	non-structural
NTC	no template control
ORF	open reading frame
PARV4	human parvovirus 4
PCR	polymerase chain reaction
qPCR	real-time quantitative polymerase chain reaction
R	correlation coefficient

RNG	range
RT-qPCR	reverse transcription polymerase chain reaction
RV	rat virus
SD	standard deviation
SIV	simian immunodeficiency virus
ssDNA	single stranded DNA
STLPyV	STL polyomavirus
TE	Tris-EDTA buffer
TSPyV	Trichodysplasia spinulosa polyomavirus
TuV	tusavirus
VP	viral protein / structural protein
WUPyV	WU polyomavirus
yr	years-old

Contents

1 Introduction

In the old days, viruses were identified when they caused severe symptoms or epidemics and pandemics among humans, domestic animals and plants. However, many viruses chronically infect hosts without causing any clinical symptoms, except perhaps in the very young, the very old, or the immunocompromised populations [1].

The improvement of sequencing and computer platforms has enabled researchers to explore the human virome – viruses resided inside and on the surface of the human body [2]. Viral metagenomic studies on different human samples have resulted in a number of novel viruses found in our gut [3, 4, 5], blood [6, 7, 8], skin [9, 10, 11] and respiratory systems [12, 13, 14]. Studying the prevalence of these novel viruses and their effects on humans can help us better decide which viruses should be prevented [15, 16], and which viruses may be ignored [8, 14], or encouraged [17, 18, 19], or engineered for clinical or biotechnological purposes [20, 21, 22].

Among these novel viruses, many are members of the parvovirus family. Parvoviruses are diminutive, nonenveloped viruses with a genome of ~ 5 kilobases and a size similar to a ribozyme. Their small size determines their parasitic strategies – they require dividing cells for their own replication, and some members of the parvovirus family (e.g. adeno-associated virus) cannot replicate without co-infection with a helper virus. The only known human pathogens in parvovirus family are parvovirus B19 and human bocaviruses. While there are many intriguing features and interesting members among this family, this thesis focuses on the detection and quantification of three most newly discovered human-associated parvoviruses – bufavirus, tusavirus, and cutavirus – all putatively belonging to *Protoparvovirus* genus.

Bufavirus identified in 2012 is the first member in *Protoparvovirus* genus found to infect humans [23]. Two more putative members of this genus, tusavirus and cutavirus, were discovered in 2014 [24] and 2016 [25], respectively. Yet, little research data can be found related to these three emerging DNA viruses. To further our understanding of bufa-, tusa-, and cutaviruses, a multiplex real-time quantitative PCR (qPCR) was developed in this study to detect and quantify all three viruses in one single reaction. The assay was then used to assess the prevalence of bufa-, tusa-, cutavirus DNA among 93 non-immunosuppressed and 137 im-

munosuppressed individuals.

The **Review of the literature** part gives a brief introduction of parvovirus, and its taxonomy and phylogeny, with an emphasis on *Parvovirinae* subfamily. Then, the literature review focuses on the three human-associated protoparvovirus, especially the epidemiology data that have accumulated so far for the three viruses.

The **Material and methods** section describes acquisition, handling and preparation of the clinical specimens used in this study. Detailed protocols of the multiplex assay and methods used for evaluation of qPCR assay are also included in this section.

In the following **Result** section, assay performance characteristics, such as PCR efficiency, analytical sensitivity, analytical specificity, repeatability, and reproducibility, are reported. The multiplex qPCR results of immunosuppressed and non-immunosuppressed cohorts are also presented.

Finally, in the **Discussion and prospects** section, the advantages and drawbacks of the multiplex assay are analyzed, followed by an interpretation of the qPCR results of clinical specimens and some future suggestions for the study of these three human-associated protoparvoviruses.

2 Review of the literature

2.1 Introduction to Parvoviruses

Parvoviruses are among the smallest known viruses, with virion diameters of 18 to 26 nanometers (nm) [26]. Hence the name “parvo”, which originates from the Latin word “parvu” meaning small. They contain a single-stranded DNA (ssDNA) genome ranging from 5 to 6 kilobases (kb) in length, which is flanked by two terminal hairpin structures [27]. Two major genes are encoded by parvoviruses, a non-structural (NS) gene that encodes replication proteins, and a viral protein (VP) gene that encodes capsid proteins.

2.1.1 History of parvoviruses

The first parvovirus, Kilham rat virus (KRV), was discovered in 1959 by Kilham and Olivier [28]. It was isolated from lysates of a rat tumor and was initially named rat virus (RV). The name “parvovirus” was first introduced by Carlos Brailovsky in a 1966 paper [29], where he denoted RV as *Parvovirus ratti*. The taxonomic family *Parvoviridae* was later adopted and approved by the International Committee on the Taxonomy of Viruses (ICTV) in the 1970s [30].

The first boost to the popularity of parvoviruses was when a pandemic occurred among puppies in 1978 caused by a canine parvovirus (CPV). A number of investigations took place showed the 1978 CPV strain (termed CPV type 2) was indeed a new strain [31] that emerged from feline parvovirus (FPV) or an FPV ancestor [32, 33].

The second boost was the discovery of the pathogenic role of human parvovirus B19 in transient aplastic crises among children with sickle cell disease [34]. B19 was first discovered in 1975 by Cossart and colleagues [35] and is the causative agent for erythema infectiosum, also known as fifth disease [36, 37]. More interests were drawn to B19 when its ability to cause miscarriage in the second trimester was shown [38].

The third boost was the potentiality of adeno-associated viruses (AAV) as the gene delivery system in gene therapy. AAVs were the first parvoviruses found to infect humans [27]. They are "defective" viruses requiring co-infection with adenoviruses or helper viruses for its replication. AAVs’ lack of pathogenicity in

humans [39, 20] has made it a tempting vector for gene therapy.

2.1.2 Taxonomy and phylogeny

The ICTV defines virus taxonomy at levels of Order, Family, Subfamily, Genus, and Species (Figure 1).

Order *(-virales)*
Family *(-viridae)*
Subfamily (*-virinae)*
Genus *(-virus)*
Species

Figure 1: The hierarchy of viral taxa defined by ICTV, with the taxon suffixes in italics.

Currently, two subfamilies of the family *Parvoviridae* have been approved by ICTV based on their host range - the *Parvovirinae*, which infect vertebrates, and the *Densovirinae*, which infect invertebrates.

The subfamilies are then further divided into genera. Within the *Parvoviridae* family, a genus is defined as monophyletic species, of which, based on pairwise alignments, the NS1 amino acid (aa) sequence share >30% identity, but <30% identity to other genera[40]. While this rule applied well to the *Parvovirinae* in separating all eight genera (Figure 2), it is applied less rigorously to *Densovirinae* subfamily due to a lack of attributable sequences [40]. There are currently five genera in the *Densovirinae* subfamily.

The lowest taxon is virus species, which is defined by the ICTV as "a polythetic class of viruses that constitute a replicating lineage and occupy a particular ecological niche" [42]. As the increasing availability of nucleotide sequence, the species taxonomy is largely determined by viral nucleotide or amino acid sequences. Within the *Parvoviridae* family, viruses in a species are required to reach >85% aa identity in the NS1 protein and >15% divergence from viruses of other species[40]. Other criteria, for instance, host, antigenic properties and genome features, are also considered for viral species classification.

The focus of this thesis is the recently discovered human-associated protoparvoviruses which belong to the *Protoparvovirus* genus. The prefix "Proto-" orig-

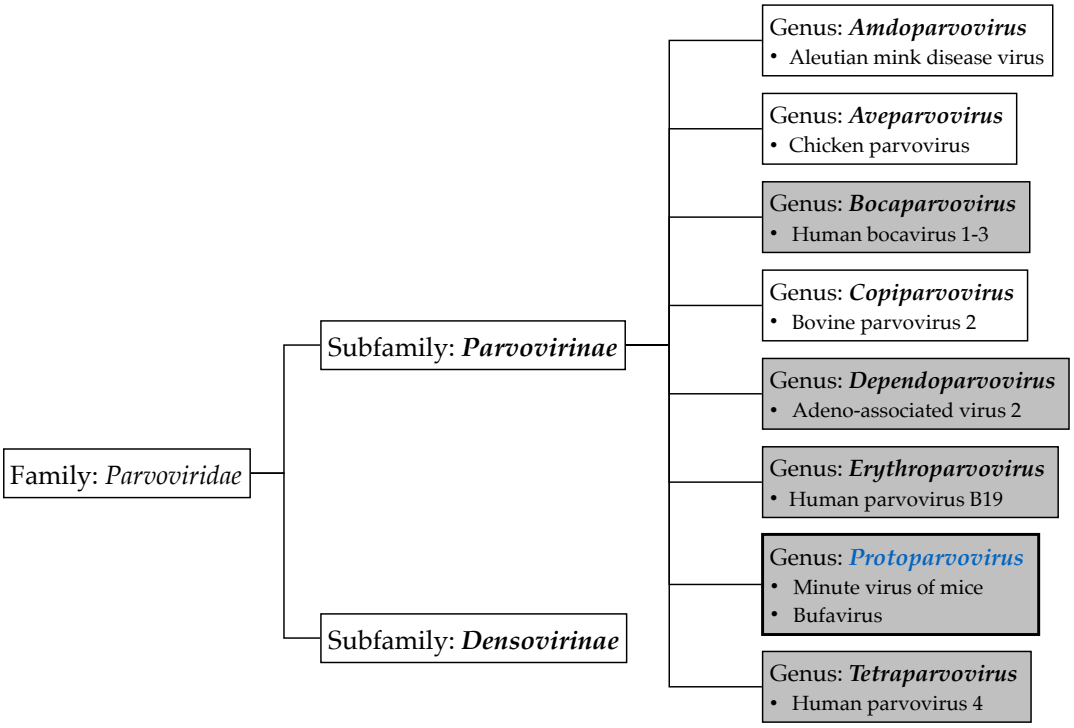

Figure 2: Taxonomy of the *Parvoviridae* family from Family to Genera based on ICTV taxonomy released in 2016 [41]. Representative viruses are listed under each genus. The gray shaded boxes indicate genera that contain human viruses. The *Protoparvovirus* genus, which is the focus of this thesis, is shown in bold frame and highlighted in blue color. Genera of *Densovirinae* are not shown here.

inated from Greek meaning "first" indicates the inclusion of the very first parvovirus, Kilham rat virus[28], to this genus. According to ICTV taxonomy report released in 2016 [41], there are currently five species within this genus, *Carnivore protoparvovirus 1*, *Primate protoparvovirus 1*, *Rodent protoparvovirus 1*, *Rodent protoparvovirus 2*, and *Ungulate protoparvovirus*, with *Primate protoparvovirus 1* species being the first virus in this genus infecting human [40]. The only official member of *Primate protoparvovirus 1* is discovered in 2012 named bufavirus [23]. Tentative members of *Protoparvovirus* genus include tusavirus discovered in 2014 [24] and cutavirus discovered in 2016 [25].

The phylogenies of official and tentative members of *Primate protoparvovirus 1*,

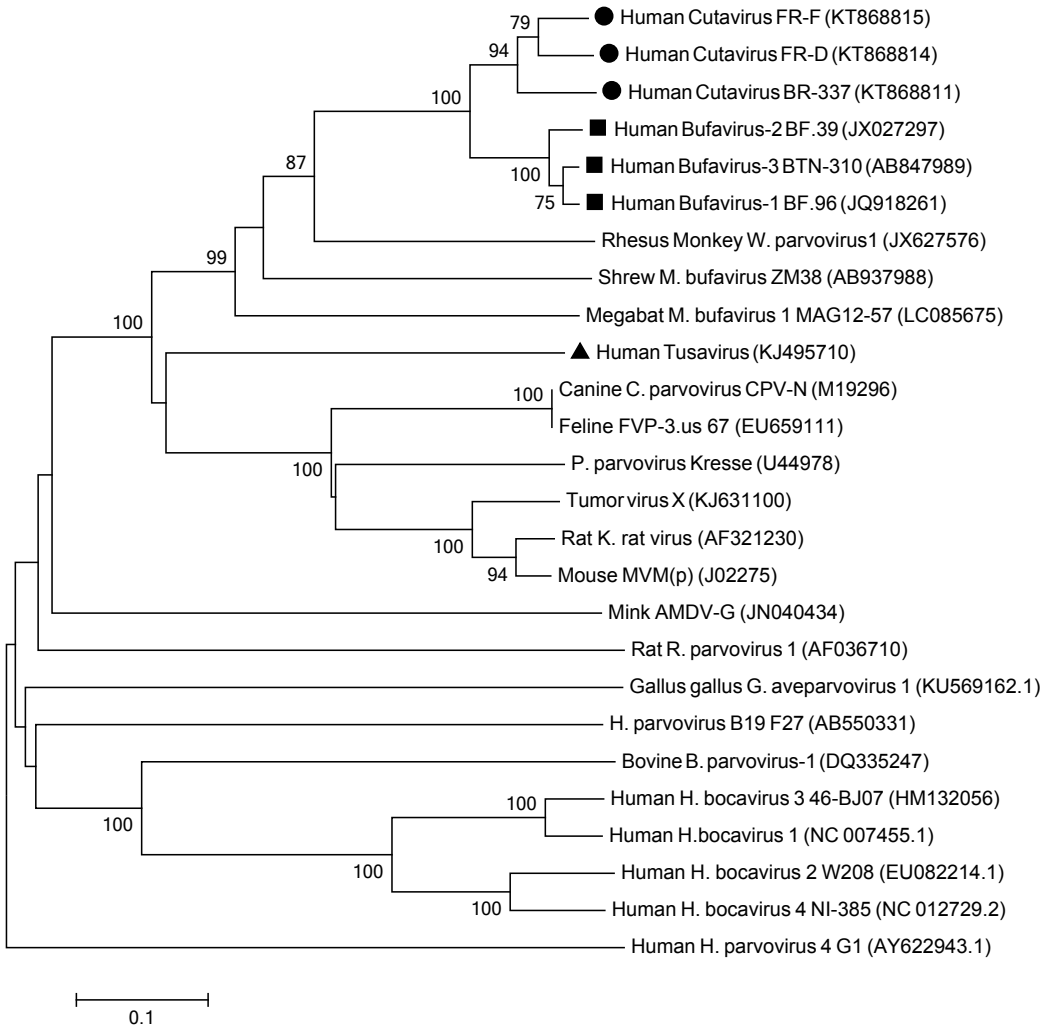

Figure 3: Phylogenetic analyses of the NS1 proteins of BuVs, TuVs, CuVs and other members of *Parvovirinae*. Sequences were aligned with Clustal W [43, 44] using PAM [45] with gap open penalty of 10 and gap extension penalty of 0.1. The phylogenetic tree was generated using the Neighbor-Joining method [46] in MEGA7 [47]. The tree is derived from 100 bootstraps [48] and bootstrap value $\geqslant$ 70% are shown. The distances were computed using the p-distance method [49] and are in the units of the number of amino acid differences per site.

other members of *Protoparvovirus* genus, and some representative members of *Parvovirinae* is shown in Figure 3 (based on NS1 protein) and Figure 4 (based on VP1 protein).

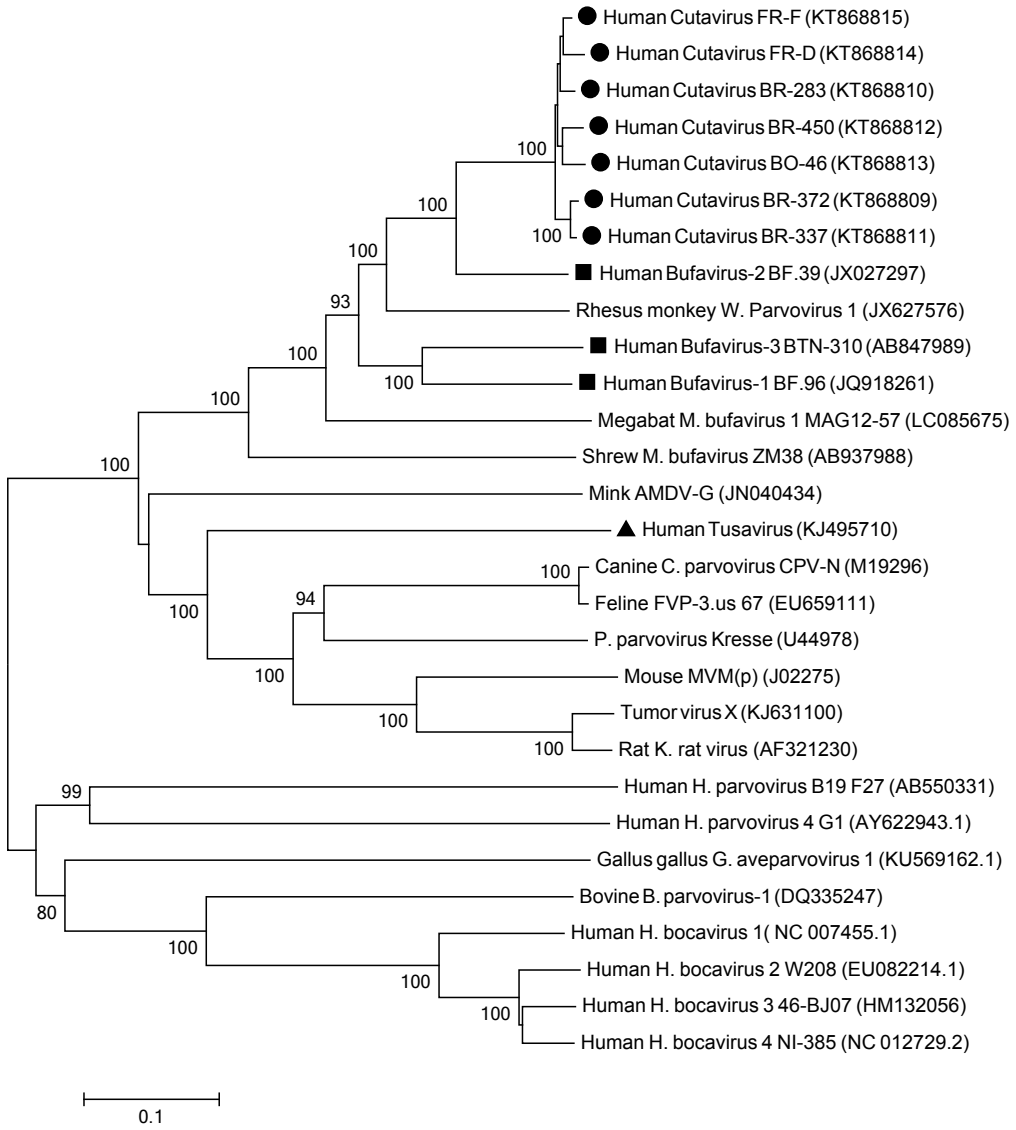

Figure 4: Phylogenetic analyses of the VP1 proteins of BuVs, TuVs, CuVs and other members of *Parvovirinae*. Sequences were aligned with Clustal W [43, 44] using PAM [45] with gap open penalty of 10 and gap extension penalty of 0.1. The phylogenetic tree was generated using the Neighbor-Joining method [46] in MEGA7 [47]. The tree is derived from 100 bootstraps [48] and bootstrap value $\geqslant$ 70% are shown. The distances were computed using the p-distance method [49] and are in the units of the number of amino acid differences per site.

As shown in Figure 3, BuVs and CuVs are phylogenetically closer to each other, whereas TuV is more distantly related to the two human-associated protoparvoviruses. The phylogeny based on VP1 (Figure 4) indicated that the VP1 regions of CuVs are more related to that of BuV genotype 2 than that of BuV genotype 1 and 3. And TuV VP1 is more distantly related to BuVs and CuVs.

2.2 Bufavirus, Tusavirus and Cutavirus

2.2.1 Discovery

Next-generation sequencing and metagenomics have revolutionized our understanding towards viruses [50]. Novel sequencing technologies [50] and sequence-independent amplification techniques [51] have accelerated the discovery of human-associated viruses and provide insights into human virome – all viruses living in and on the human body [2, 52, 53].

A series of parvoviruses discoveries were made by Dr. Delwart and colleagues [54, 55, 56, 23, 24, 25]. Such viral metagenomics approach to identify novel viruses includes five steps [57]: 1) enrichment and purification of viral nucleic acid using filtration and enzymatic digestion; 2) random amplification of extracted DNA through sequence independent approach such as random PCR; 3) subclone amplified products; 4) sequencing of the clones; 5) sequence alignment to known virus sequences in public databases.

Bufavirus Human Bufavirus (BuV), the founding and only official member of *Primate protoparvovirus 1* species, was discovered in 2012 by a metagenomics study on 98 children (< 5yr) with acute diarrhea in Burkina Faso[23]. As a result, sequences from one specimen showed significant similarities to parvovirus proteins. NS1 and VP1 aa identities were 39% and 31% to the closest parvovirus members. PCR screening of the 98 fecal samples resulted in a total 4/98 (4%) BuV DNA positive. Sequencing of the PCR amplicons resulted in BuV 1 and BuV 2 genotypes, which highly diverse in the capsid gene (72% identity in VP1). BuV was initially proposed as a new *Parvoviridae* genus, and later, in 2013, was approved to be the founding member of *Primate protoparvovirus 1* species under *Protoparvovirus* genus [40].

In 2014, BuV 3 genotype was found in three fecal samples from children with

diarrhea in Bhutan (3/393 or 0.8%) [58]. NS1 of BuV 3 showed 95 - 96% aa identity to that of BuV 1 and BuV 2. As for VP1, aa sequence of BuV 3 shared 73% with BuV1 and 71% with BuV 2.

Tusavirus In 2014, another candidate virus for *Protoparvovirus*, called tusavirus (TuV), was identified in the fecal sample of an 18 month-old girl with unexplained diarrhea in Tunisia (0.56% or 1/180) [24]. No other mammalian viruses were detected in this sample. TuV NS1 and VP1 showed closest identity of 44% and 39% identity to those of Kilham rat parvovirus [24]. Thus, TuV was purposed to be a new member of the *Protoparvovirus* genus and to be the prototype species of *primate protoparvovirus 2*. However, whether TuV is truly a human parvovirus requires further investigation.

Cutavirus In 2016, another putative *Protoparvovirus*, cutavirus (CuV), was detected in diarrhea samples from Brazil (1.6% or 4/245) and Botswana (1% or 1/100) using viral metagenomics [25]. The genome of CuV showed 76% and 82% identity to NS1 and VP1 aa sequence of the closest genome, human BuV2, respectively, suggesting CuV may be a distinct protoparvovirus species. A digital and PCR screening resulted in the detection of CuV genome from two French skin biopsies with cutaneous T cell lymphomas (CTCL). A further PCR assay resulted in a total number of 4/17 (23.5%) CuV DNA positive CTCL skin tissue samples. The role of CuV in diarrhea and CTCL remains to be further investigated.

2.2.2 Genome

Genomes of parvovirus are typically linear ssDNA molecule 5 to 6 kb in length [27], with short imperfect terminal palindromes that flank the coding sequence folding back to form hairpin telomeres [30].The hairpins are served as primers for complementary strand synthesis in its unique rolling hairpin replication [30, 59]. A typical protoparvovirus genome is shown in Figure 5. Two major genes are encoded by parvoviruses – a non-structural (NS) gene in 5′ end region and a structural (VP) gene in the 3′ end region.

The genome for the three human-associated protoparvoviruses, BuV, TuV, and CuV, are $\sim$ 5 kb in length encodes three major proteins, NS1, VP1 and VP2, as shown in Figure 6. The 5′ ORF encodes the NS1 protein, while the 3′ORF encodes

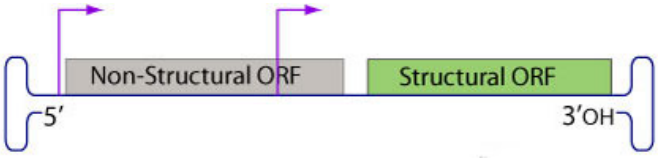

Figure 5: Protoparvovirus genome adapted from ViralZone [60]. Hairpin structures are illustrated at the termini. There are two open reading frames (ORFs) for non-structural and structural proteins, respectively.

the two structural proteins, VP1 and VP2. VP1 and VP2 are translated from the same transcripts using alternative splicing, and they differ in the N-termini. The phospholiapse A_2 motif was identified in VP1 ORFs of all three viruses. Both BuV and CuV genomes contain a theoretical middle ORF with unknown function and showed no similarity to any parvovirus proteins or to each other.

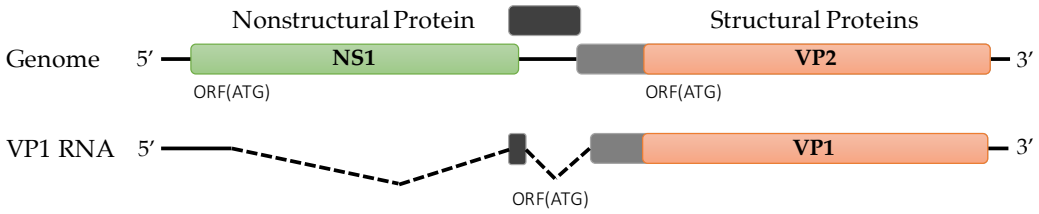

Figure 6: Organization of human associated protoparvovirus, adapted from [23, 24, 25]. Putative ORFs are shown for NS1, VP1 and VP2. Genomes are transcribed into mRNA utilizing host proteins. Alternative splicing of mRNA allows the expression of structural proteins.

2.2.3 Epidemiology

Bufavirus The search for human BuVs was largely conducted in stool samples among patients with diarrhea or gastroenteritis. To date, it has been detected in diarrhea samples of children in Burkina Faso [23], Tunisia [23, 61], Bhutan [58], Thailand [62], Turkey [63], China [64], and Finland [65], as well as in diarrhea samples of adults in Finland [66], the Netherlands [67], Thailand, and China. BuV DNA was also detected in nasal swab sample in Finnish children population with a prevalence of 0.1% (1/955) [65]. The list of studies assess the BuV DNA

prevalence are summarized in Table 1.

Study	Specimen	Country	Age (years)	Positive (%)	n	Symptom	Genotype
[23]	stool	Burkina Faso	median NA RNG 0-5	3.1	98	GE	BuV1
[23]	stool	Tunisia	"children"	1.6	63	NPAFP	BuV1
[62]	stool	Thailand	median NA RNG 0-97	0.3	1495	GE	BuV1
[64]	stool	China	median NA RNG 0-85	0.2	1877	GE	BuV1
[61]	stool	Tunisia	median 0.6 RNG 0-5	1.0	203	GE	BuV1
[66]	stool	Finland	median 51.5 RNG 0-99	1.1	629	GE	BuV1
[23]	stool	Burkina Faso	median NA RNG 0-5	1.0	98	GE	BuV2
[58]	stool	Bhutan	median NA RNG 0-5	0.8	393	GE	BuV3
[64]	stool	China	median NA RNG 0-85	0.3	1877	GE	BuV3
[63]	stool	Turkey	AVG 1.63 RNG 1-5	0.7	583	GE	BuV3
[67]	stool	the Netherlands	median 47 RNG 0-97	3.7	27	GE	BuV3
[63]	stool	Turkey	AVG 1.63 RNG 1-5	0.7	583	GE	unknown
[65]	stool	Finland	median 1.17	1.2	172	GE	unknown
			RNG 0-15.6	0	545	ARTI	
				0.4	238	GE & ARTI	
[65]	nasal swab	Finland	median 1.17	0	172	GE	unknown
			RNG 0-15.6	0	545	ARTI	
				0.4	238	GE & ARTI	
[23]	stool	Chile	"children"	0	100	GE	-
[62]	stool	Thailand	median NA RNG 0-39	0	726	HFMD	-
[64]	stool	China	"children"	0	421	non-GE	-
[63]	stool	Turkey	AVG 1.44	0	148	healthy	-

Table 1: Genotype stratified BuV DNA prevalences in stool or nasal swab samples of subjects with or without diarrhea.
ARTI, acute respiratory tract infection; AVG, average; GE, gastroenteritis; HFMD, hand-foot-and-mouth disease; NA, not available; NPAFP, non-polio acute flaccid paralysis; RNG, range

The BuV1 DNA has been detected in fecal samples of all age groups in Africa, Europe and Asia; the BuV2 DNA has only been detected in Africa; and the BuV3 DNA has predominately been found in Asia. The prevalence of BuV in symp-

tomatic patients has been low, ranging 0.3-3.7%. Though four studies [65, 62, 64, 63] were conducted on non-diarrhea patients or healthy individuals, the results of which were all BuV DNA negative, the role of the virus in diarrhea was still uncertain. Moreover, one study conducted on non-polio acute flaccid paralysis [23] resulted in one BuV1 positive, though the diarrhea condition of the patient was not reported in the original study.

A serological study was conducted to assess BuV seroprevalence [65] in 180 healthy adults and 228 children with gastroenteritis and/or acute respiratory tract infection. The results of BuV IgG seropositive subjects are listed in Table 2, with 10/180 (5.6%) in adults and 7/228 (3.1%) in children. Of the 180 adults, 163 were from Finland, whereas 12 were Asian origin, 1 from America and 4 from other European countries. The results suggested BuVs infection may be more prevalent in Asia with 6/12 (50%) seropositive.

Genotype	Adults				Children
	Middle East	India	China	Finland	Finland
BuV1	1[a]	3			2[b]
BuV2	1[a]		1	5	4[b]
BuV3	1[a]				2
BuVs	1	3	1	5	7

Table 2: BuV IgG positive adults and children. The origin of adults are shown in the header. The "BuVs" row sums the overall numbers of individuals who are BuVs IgG positive.

[a] One individual was positive for all three BuVs.

[b] One individual was positive for both BuV1 and BuV2

The first BuV-related sequences detected in non-human primates were named WUHARV parvoviruses, which were most closely related to BuV2 [23, 68]. They were obtained form fecal and serum samples of rhesus monkeys (*Macaca mulatta*) experimentally infected with simian immunodeficiency viruses (SIVs) in the United States [68]. BuV-related sequences were also identified in spleen tissues from wild yellow baboons(*Papio cynocephalus*) and a chacma baboon (*Papio ursinus*) from Zambia [69]. BuV-related sequences have also reported in fecal samples of Hungarian microbats (*Miniopterus schreibersii*) [70], pharyngeal or anal swab of microbats in China [71], feces, spleen and liver samples of wild shrews of the *Crocidura* genus in Zambia [69], feces of wild rat in China [72], feces of domestic pigs from Hungary [73, 74] and Austria [74], spleen and fecal samples of mega-

bats in Indonesia [75], as well as feces of fur seals in the Subantarctic and South America [76].

Tusavirus So far, the detection of TuV DNA has only been described in the original discovery study, where it was presented in the fecal sample of one child with unexplained diarrhea in Tunisia (0.56% or 1/180) [24]. Serologic study on 228 Finnish children has reported one child as TuV-IgG positive (0.4%) [65]. TuV-related sequences were found in fecal samples of Subantarctic and South America fur seals [76]. More studies are required to confirm whether TuV is a human virus.

Cutavirus In the original discovery study [25], CuV DNA was found in fecal samples of children with diarrhea in Brazil (1.6% or 4/245) and individuals with diarrhea in Botswana (1.0% or 1/100), as well as skin samples of patients with mycosis fungoides, the most common type CTCL, in France (23.5% or 4/17). 31 non-CTCL skin samples were PCR screened and negative for CuV DNA. One Danish patient with cutaneous malignant melanoma (10% or 1/10) was later found CuV DNA positive in lesion area [77]. A detailed list of prevalence of CuV DNA in human samples can be found in Table 3. These findings indicate that the presence of CuV DNA is not limited to fecal samples but can also be observed in skin tissues. Nevertheless, further evidences will be needed to resolve its role in enteric diseases and cutaneous cancers.

Study	Specimen	Country	Positive (%)	n	Symptom
[25]	stool	Brazil	1.6	245	diarrhea
[25]	stool	Botswana	1.0	100	diarrhea
[25]	skin	France	23.5	17	mycosis fungoides
[77]	skin	Denmark	10	10	cutaneous malignant melanoma
[25]	skin	France	0	10	skin carcinoma
[25]	skin	4 France; 6 ND	0	10	parapsoriasis
[25]	skin	ND	0	4	eczema
[25]	skin	ND	0	4	eczematoid dermatitis
[25]	skin	ND	0	3	healthy

Table 3: CuV DNA prevalences in stool samples of patients with diarrhea and skin biopsy samples of patients with different skin diseases. ND, no data.

3 Aims of the study

There were two main goals of the study

- to develop and evaluate a real-time multiplex quantitative PCR (qPCR) assay for detection of BuV, TuV, and CuV DNA
- to assess the prevalence of BuV, TuV, CuV DNA among Finnish population

4 Material and methods

4.1 Specimens

4.1.1 Immunosuppressed cohort

Skin biopsies were obtained from 137 immunosuppressed transplant recipients, consisting of 131 Finnish liver-transplant patients, four Finnish kidney-transplant patients and one heart-transplant patient. All of the patients were above 18 years old and accepted transplant more than five years ago. 134 serum samples were also acquired from the liver-transplant patients cohort. However, these serum samples were not completely paired with the skin tissue samples.

For PCR, DNA from skin biopsies ($\sim$ 4 to 5 mm diameter) and serum (200 μL) was extracted with the DNA Blood Mini Kit (Qiagen) according to the manufacturer's instruction for tissue DNA and blood DNA, respectively. The extracted DNA was eluted in 100 μL AE buffer (kit provided). All samples were stored at -80 °C, and all DNA extracts were stored at -20 °C.

4.1.2 Non-immunosuppressed cohort

144 extracted DNA from 93 non-immunosuppressed individuals with allergic or irritant contact dermatitis was also included as control for the purpose of this study. The median age was 43 years old (range 18 to 67 years). Of the 93 individuals, 42 were diagnosed with allergic contact dermatitis (ACD) and 51 were diagnosed with irritant contact dermatitis (ICD). The skin samples were taken as a 4 mm punch biopsy from the test reactions on the back of the patient. Samples were stored at -80 °C, and DNA extracts were stored at -20 °C.

The Ethics Committee of the Hospital District of Helsinki and Uusimaa approved the study, and the study was conducted in accordance with the relevant guidelines and regulations.

4.1.3 BuV and TuV DNA positive samples

BuV and TuV DNA positive samples were generous gifts from Dr. Eric Delwart. Four BuV positive (BuV1-BF7, BuV1-BF86, BuV1-BF96, BuV2-BF39) and one TuV

positive fecal suppressant were from the original metagenomics studies, where the viruses were discovered [23, 24]. DNA extraction was performed with QIAamp DNA mini kit (Qiagen) using 100 μL of fecal supernatant in 2012 for BuV positive samples, and in 2015 for TuV positive sample. All DNA extracts have then been stored at -20 °C.

4.2 BuV, TuV, CuV multiplex qPCR assay

Singleplex BuV qPCR was developed previously by Väisänen et al. [66]. In this study, individual singleplex TuV and CuV qPCR reactions were developed and validated to test specificities of primer sets and probes. Each assay was run in singleplex reactions to test the assay efficiency, sensitivity and specificity.

After setting up individual singleplex qPCRs, single reactions were combined into one multiplex qPCR for detection and quantification of all three human-associated protoparvoviruses. The performance of multiplex qPCR was compared to that of the singleplex qPCRs to verify that quantification cycle (C_q) obtained from multiplex were not artificially reduced by multiplex reaction.

4.2.1 The principle of assay

Detection chemistries for qPCR To detect and quantify the amount of target, a measurable signal must be generated in the courses of PCR and is proportional to the amount of amplified product. All current detection methods utilizes fluorescent technology. There are now two main types of detection methods: non-specific techniques based on DNA binding agents and specific techniques based on fluorescence resonance energy transfer (FRET). For this study, SYBR® Green and hydrolysis probes were used for detecting fluorescent signals in qPCR.

As shown in Figure 7, SYBR® Green is an intercalating dye that binds to double-stranded DNA in the reaction and fluoresces once it is bound. The non-specific nature of SYBR® Green means that it not only binds to the amplified products, but also to primer-dimers and other unspecific products. However, since each product has its own melting temperature depending on its size and base contents, the specificity of the product can be verified by a melting curve analysis after the PCR amplification. Therefore, in this study, SYBR® Green was used to check the specificity of primers, to detect undesired PCR products, and to help

the optimization of the qPCR system.

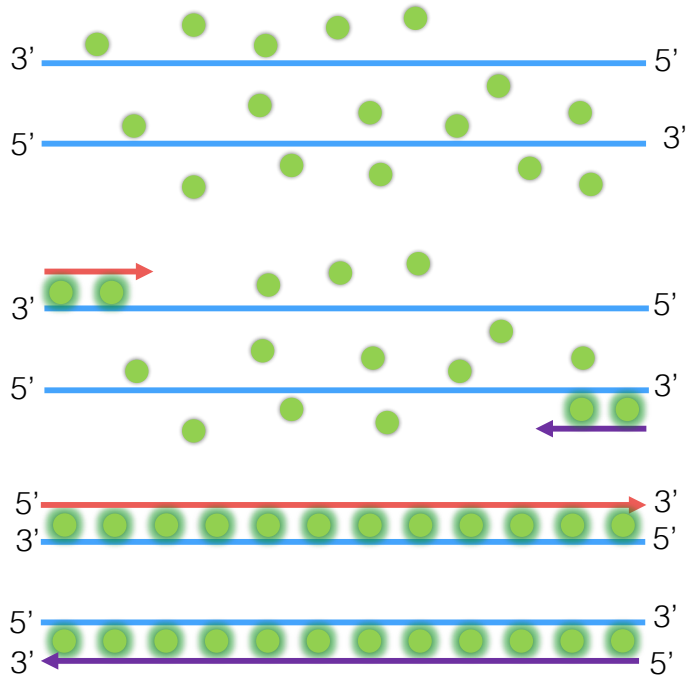

Figure 7: SYBR® Green Chemistry. DNA is denatured as ssDNA and SYBR® Green molecules are free in the reaction and do not emit fluorescence. SYBR® Green binds to dsDNA as primer annealing and polymerase elongation take place. The fluorescence signal increase exponentially and is proportional to dsDNA presented in the reaction.

Hydrolysis probes, or TaqMan® probes, consist of a single-stranded probe sequence, a fluorophore attached to the the 5′ end of the probe, and a quencher attached to the 3′ end of the probe. The principle of hydrolysis probe based quantification is shown in Figure 8. When the fluorophore is excited by light at a specific wavelength, it passes its energy to the quencher via FRET and thus the energy is dissipated in the form of light or heat instead of fluorescence. As long as the fluorophore and the quencher are in close proximity, very little fluorescent signal can be detected. During the annealing step, the probe binds to the amplicon; when DNA polymerase extends the primer, it hydrolyses the probe due to its 5′-3′ exonuclease activity and releases the fluorophore and quencher into the solution. Thus, the emission of fluorophore is no longer quenched and an increase

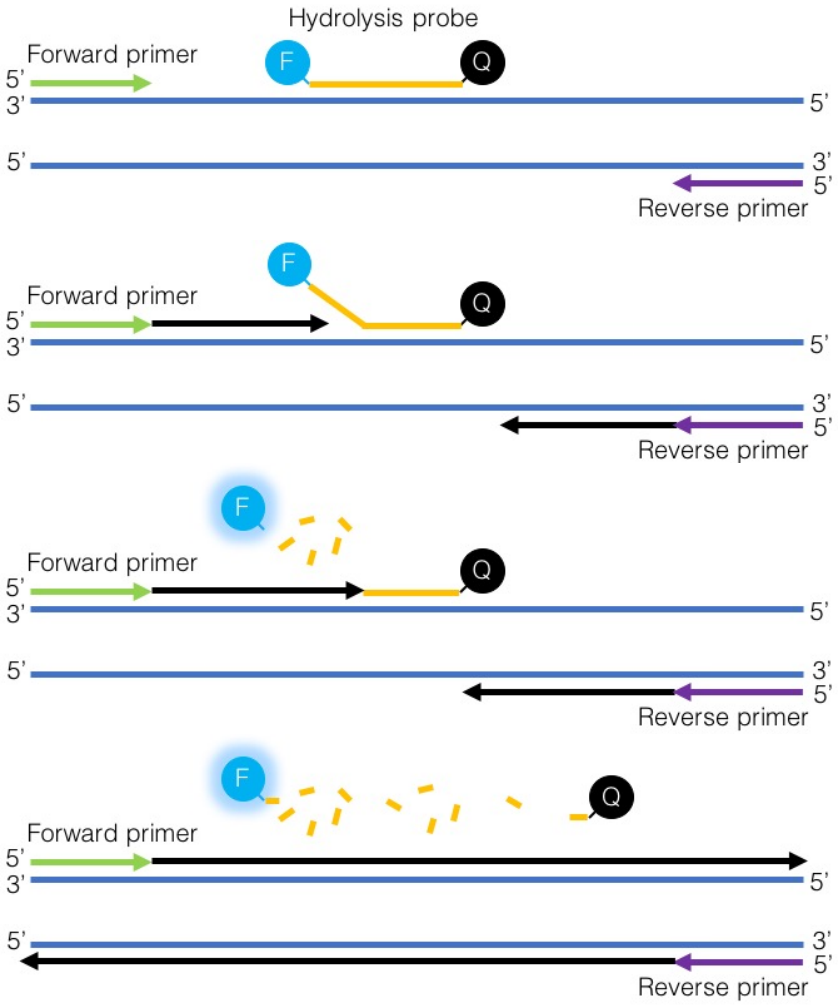

Figure 8: Hydrolysis probe chemistry. Primers and probe anneal to target sequence. The quencher (Q) prevents fluorescence emission when quencher and fluorophore (F) are in close proximity. When DNA polymerase begins to extend the primer and reaches the probe, the probe is hydrolysed and fluorophore is no longer quenched and emits fluorescence signal. The fluorescence signal is proportional to the abundance of amplified targets.

in fluorescence can be measured. Hydrolysis probe offers better specificity, since a positive result requires specific binding and the probe is in the correct orientation. By utilizing different fluorophores to label different probes, the fluorescent signals generated by several products can be simultaneously detected.

Quantification Two common quantification strategies are used for real-time PCR – absolute quantification and relative quantification. Absolute quantification requires serially diluted samples with known concentration or copy number to generate standard curve. Relative quantification, is used to determine the relative change in target molecule in a sample versus in a reference sample, thus a dilution series with unknown concentration can be used. Absolute quantification is usually used in qPCR assay for pathogen detection, whereas relative quantification is usually used in reverse transcription polymerase chain reaction (RT-qPCR). For the purpose of this study, absolute quantification strategy was used.

Quantification cycle value, C_q value, represents number of cycles needed to cross the threshold signal level. In absolute quantification, a dilution series, which contains the same sequence as the target, with known concentration or copy number is used to generate a standard curve. The standard curve is a plot of C_q value versus the logarithm of the standards' number of template copies. Based on the standard curve, the amount of the target in a sample can be determined with its C_q value.

4.2.2 Plasmid controls

The multiplex qPCR for BuV, TuV and CuV requires plasmid controls that contain part of the viral genome for validation and building the standard curve. For this purpose, a 3.5 kb piece of BuV1 NS1 region (16-3577 nt, JX027295) amplified from the original fecal supernatant, a $\sim$ 1.7 kb piece of TuV VP2 region(2659 - 4356 nt, KJ495710.1), and a $\sim$ 1.7 kb piece of CuV VP2 region (2747 - 4456 nt, KT868811) amplified from CuV-BR-337 from a Brazilian diarrheic stool sample, were cloned into a pSTBlue vector.

All plasmids were purified using QIAprep Spin Miniprep Kit (Qiagen) and their concentrations were measured by NanoDrop 1000 Spectrophotometer (Finnzymes) and then converted to the number of copies using the molecular weight of the DNA. Afterwards, the plasmids were diluted into a 10-fold dilution series from

10^0 to 10^7 copies/μL with TE buffer (10mM Tris: 0.1mM EDTA; pH 8.0), and stored at -20 °C. These plasmids were used to test the sensitivity and specificity of both singleplex and multiplex qPCRs.

4.2.3 Primers and hydrolysis probes

Primers and hydrolysis probes were designed for both TuV and CuV, but for BuV a working singleplex qPCR was already available [66]. Primers and probes were designed to give maximum intraspecies similarity and no interspecies similarity. The conversed and variable regions of BuV1-3, TuV and CuVs were identified by aligning 14 BuV1-3 sequences, the one existing TuV sequence, and six CuV sequences using Clustal Omega (`http://www.ebi.ac.uk/Tools/msa/clustalo/`) [78]. The GenBank accession numbers and geographical origins of the sequences are listed in Table 4.

For the BuV assay, the primer region is within the NS1 gene [66]; but for TuV and CuV, the primer regions were designed within the VP2 gene. In order to be used together as a multiplex, the NetPrimer online software (`http://www.primerbiosoft.com/netprimer/index.html`) was used to identify any complementarities between primers and probes for all three viruses. No significant secondary structures were detected (data not shown). Amplify4 (`http://engels.genetics.wisc.edu/amplify/`) was used to simulate and evaluate the multiplex qPCR set-up. Blasting with NCBI's Basic Local Alignment Search Tool (`https://blast.ncbi.nlm.nih.gov/Blast.cgi`) [79] of primers and the probe against GenBank sequences was performed to identify any undesired matching sequences.

In this study, an unique hydrolysis probe was designed for each virus. The selection of fluorophores was based on three principles: (a) the emission wavelengths of the fluorophores should be detectable by the qPCR instrument; (b) the fluorophores should have little to no overlap in their emission spectra to avoid overlap of the signal; (c) the signal intensities should also be high enough to allow detection of low copy-number targets. Therefore, many trails were conducted for the selection of the optimal combination of dyes.

The primer and probe sequences are listed in Table 5. One degenerate primer (CuV fwd) and two degenerate probes (BuV NS1 probe and CuV VP2 probe) were designed to cover all known genotypes of BuV and CuV. All primers and

Virus	Geographical origin	Accession numbers	Reference
BuV1-BF7	Burkina Faso	JX027295	[23]
BuV1-BF86	Burkina Faso	JX027296	[23]
BuV1-BF96	Burkina Faso	JQ918261	[23]
BuV1-BJ133	China	KM580347	[64]
BuV1-BJ181	China	KM580351	[64]
BuV2-BF39	Burkina Faso	JX027297	[23]
BuV3-BJ154	China	KM580348	[64]
BuV3-BTN63	Bhutan	AB847987	[58]
BuV3-BTN109	Bhutan	AB847988	[58]
BuV3-BTN310	Bhutan	AB847989	[58]
BuV3-AHP178	Turkey	AB982217	[63]
BuV3-AHP368	Turkey	AB982220	[63]
BuV3-AHP740	Turkey	AB982222	[63]
BuV3-AHP692	Turkey	AB982221	[63]
TuV	Tunisia	KJ495710	[24]
CuV-BR283	Brazil	KT868810	[25]
CuV-BR337	Brazil	KT868811	[25]
CuV-BR372	Brazil	KT868809	[25]
CuV-BR450	Brazil	KT868812	[25]
CuV-BO46	Botswana	KT868813	[25]
CuV-FR-D	France	KT868814	[25]
CuV-FR-F	France	KT868815	[25]

Table 4: Geographical origins and GenBank accession numbers of sequences used to construct alignment for the design of primers and probes.

probes were manufactured by Sigma-Aldrich® and diluted in TE buffer (10mM Tris: 0.1mM EDTA; pH 8.0).

4.2.4 Real-time quantitative PCR

All quantitative PCR were done using AriaMx Realtime PCR System (Agilent Technologies). To avoid contamination, the handling of samples, master mix components, and plasmid templates were performed in three separate rooms and all templates were added in laminar hoods. Aerosol resistant filter tips and disposable clear polypropylene qPCR 8-tube strips (4ti-0753/c, 4titude®) or 96-well

Oligo name	Sequence(5'to 3')	Amplicon size (bp)
BuV fwd [66]	ACAGTGTAGACAGTGGATTCAAACTT	
BuV rev [66]	GTTGTGGTTGGATTGTGGTTAGTTC	126
BuV NS1 probe [66]	FAM-CGGAAGAGATTTTGACAGTGCYTAGCAA-BHQ1	
TuV fwd	CCAGAAAGCCGTATCACCAT	
TuV rev	AACCAAGTGTTTCTGATCTTATTGCT	118
TuV VP2 probe	TxRd-ACACCAACAATCAACTGCCATACACACC-BHQ2	
CuV fwd	TAACACATCCCAGAATYGTCACATA	
CuV rev	TTCCATTGTCTTGGAGTGCG	91
CuV VP2 probe	JOE-AGTTKTCCTGACCACCAGAAGGTTCCA-BHQ1	

Table 5: Primers and probes used for BuV, TuV, CuV qPCR multiplex assay. "fwd" means forward (sense) primer while "rev" means reverse (antisense) primer. "NS1" means amplicon is in the NS1 region while "VP2" means amplicon is in the VP2 region. PCR amplicon sizes are shown. The International Union of Pure and Applied Chemistry (IUPAC) system is used from nucleotide nomenclature: Y = C + T; K = G + T.

plates (4ti-0740, 4titude®) were used. All runs included water as no template controls (NTC). Flat optical caps (4ti-0751, 4titude®) were used for the sealing of qPCR system.

SYBR® Green based qPCR system The SYBR® Green based qPCR assay consists of 1x Maxima SYBR Green qPCR Master Mix (Thermo Scientific) with 10 nM ROX as passive reference dye, 0.5 μM concentration of primers except for CuV fwd primer with 1.0 μM, 5 μL template, and molecular biology-grade water up to a total volume of 25 μL. Only one primer pair was used in singleplex reactions whereas all primers were included in the multiplex reaction.

After initial denaturation and enzyme activation at 95°C for 10 min, 45 cycles of amplification (95°C for 15 s, 62°C for 1 min) were performed. Melting curve analysis was then performed on PCR products using a dissociation protocol comprising 95°C for 1 min, followed by incremental temperature starting from 40 to 95°C at 0.5°C/s melt rates with Savitzky-Golay curve-smoothing setting averaging nine points. All runs include plasmid and NTC controls. The baseline fluorescences were determined by the Aria Software Version 1.3 (Agilent Technologies) baseline algorithm. The background based threshold was calculated as ten times the standard deviation of the mean baseline fluorescence in cycles 3 through 6.

Hydrolysis probe based qPCR system The qPCR reaction for hydrolysis probe assay consists of 1x Maxima probe qPCR Master Mix(Thermo Scientific) without ROX as passive reference dye, 0.5 μM concentration of primers except for CuV fwd primer with 1.0 μM, 0.2μM BuV and TuV probes, 0.4 μM CuV probe, 5 μL template and molecular biology-grade H_2O up to a total volume of 25 μL. Only one primer pair and the corresponding probe was used in singleplex reactions whereas all primers and probes were included in the multiplex reaction.

After initial denaturation and enzyme activation at 95°C for 10 min, 45 cycles of amplification (95°C for 15 s, 62°C for 1 min) were performed. All runs include plasmid and NTC controls. And NTC controls have remained negative throughout the study. The baseline fluorescences for BuV and TuV assay were determined by the Aria Software Version 1.3 (Agilent Technologies) baseline algorithm. The cutoff for BuV and TuV cycle determination was calculated as ten times the standard deviation of the mean baseline fluorescence in cycles 5 through 9, whereas the cutoff for CuV cycle determination was calculated as ten times the standard deviation of the mean baseline fluorescence in cycles 8 - 11 due to the high level of initial noise.

4.2.5 Analytical sensitivity

Analytical sensitivity is expressed as the limit of detection (LOD), which is defined as the concentration at which 95% of the positive samples are detected [80]. The sensitivity of the qPCR assay was determined by using serially diluted control plasmids with or without 90 ng human DNA per reaction. A generalized linear model was fit to the data to determine the plasmid copy number that yield 95% probability of a positive result.

4.2.6 Analytical specificity

Analytical specificity refers to the qPCR assay detecting the target sequence rather than other, nonspecific targets also present in a sample [80]. The specificity of singleplex and multiplex qPCR assay was evaluated with 90 ng human DNA per reaction from cultured HEK293, HeLa, hFSF, Hep G2, and A549 cells. DNA from these cells was extracted with the DNA Blood Mini Kit (Qiagen) according to the manufacturer's instruction for cultured cell DNA extraction, except that DNA was eluted into 100 μL AE buffer to yield a high DNA concentration. The ex-

tracted DNA was stored at -20 °C.

The specificity was also tested with 5 x 10^6 copies per reaction cloned full or near full length genomes of parvovirus B19 genotype 1, human bocavirus (HBoV) 1, 2, 3, and 4, human parvovirus 4 (PARV4), as well as a mix of 13 human polyomaviruses (HPyV), including Merkel cell polyomavirus (MCPyV), Trichodysplasia spinulosa polyomavirus (TSPyV), HPyV9, HPyV12, New Jersey polyomavirus (NJPyV), BK polyomavirus (BKPyV), JC polyomavirus (JCPyV), KI polyomavirus (KIPyV), WU polyomavirus (WUPyV), HPyV6, HPyV7, MW polyomavirus (MWPyV), and STL polyomavirus (STLPyV).

4.2.7 Repeatability and reproducibility

Variation in qPCR results can be the results of stochastic variation and differences in temperature and concentration, the latter of which is usually introduced by pipetting errors. The precision of qPCR is most commonly due to concentration variation and decreases with copy number [80]. Repeatability or intraassay variance refers to the precision and robustness of the assay with the same samples repeatedly analysed in the same run. Reproducibility or interassay variance refers to the variation of the singleplex and multiplex assays were studied using three replicates each with a ten-fold dilution series.

4.2.8 *RNase P* qPCR

Each skin tissue sample was subjected to the reference gene, *RNase P* gene, qPCR for cell count quantification. Viral quantities were expressed as copies per 10^6 cells. The qPCR was performed by previously described *RNase P* primers and probe [81, 82]. The PCRs reaction consists of 1x TaqMan® Universal PCR master mix (Applied Biosystems) with ROX as passive reference dye, 0.3 μM concentrations of forward and reverse primers, 0.15 μM probe, 5 μL of template, and molecular biology-grade water for a final volume of 25 μL. AmpliTaq Gold polymerase activation and initial denaturation was set at 95°C for 10 min. The application step consists of 40 cycles of 45 s at 95°C, 1 min at 60°C and 1 min at 72°C.

5 Results

5.1 Hydrolysis probe chemistries for the multiplex qPCR

In hydrolysis probe based qPCR assay, probes for three viruses were labelled with different report dyes: BuV with FAM (emission wavelength $\sim$ 520 nm), CuV with JOE (emission wavelength $\sim$ 552 nm) and TuV with Texas Red (emission wavelength $\sim$ 615 nm). The emission spectra of these three dyes are shown in Figure 9. The emission spectra are fairly separated and no bleed-through was observed in multiplex qPCR assay. The significant overlap between emission spectra of Texas Red and that of ROX (emission wavelength $\sim$ 605 nm) prevented the use of ROX as a common passive reference dye for fluorescence signal normalization. However, the use of a passive reference dye is optional for modern qPCR systems such as AriaMx, since well position effects are eliminated by employing an even illumination and detection system for each well instead of using one light source and detector for a whole plate.

5.2 Amplification efficiency

The efficiency of a qPCR assay is related to the slope of the standard curve. The equation is shown below:

$$PCR\ efficiency = \left((10^{-1 \times slope}) - 1\right) \times 100\%$$

Figure 10 shows the results of qPCR quantification of dilution series from 10^1 to 10^6 copies/μL of BuV, TuV, and CuV plasmids along with standard curve derived from each assay. The efficiencies ranged from 92.64% to 105.63% for singleplex assays, and from 100.53% to 103.76% for multiplex assays (Table 6). The standard curves for the singleplex assays and that for the multiplex assays are almost indistinguishable (Figure 10).

5.3 Sensitivity

Sensitivities were first determined with serial dilutions of control plasmids. At 10 copies/μL, 100% of 20 replicates were positive with both multiplex and singleplex assays with or without 90 ng human DNA, which indicates a LOD of less

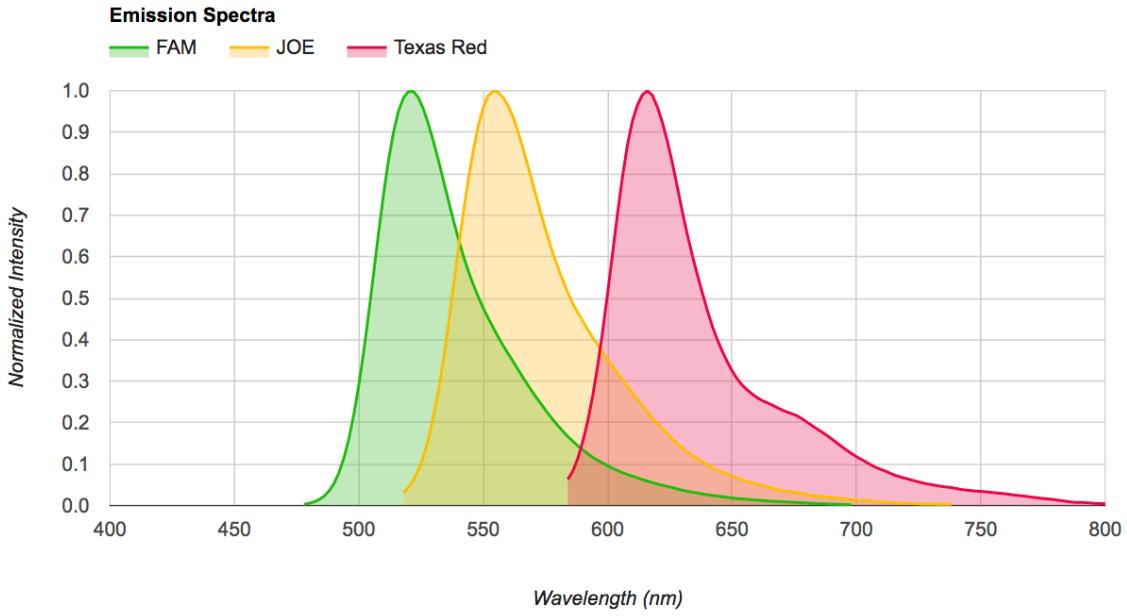

Figure 9: Emission spectra of fluorophore dyes – FAM, JOE and Texas Red. X axis denotes the wavelength (nm). Y axis denotes the normalized fluorescence intensity. Pictured was derived from LGC Biosearch Technologies spectral overlay tool for multiplex qPCR (`http://www.qpcrdesign.com/spectral-overlay`).

copies/μL	Intraassay variation (Mean C_q value $\pm$ SD)					
	Singleplex			Multiplex		
	BuV	TuV	CuV	BuV	TuV	CuV
Efficiency	92.64%	103.78%	105.63%	100.53%	103.25%	103.76%
R^2	0.998	0.998	0.997	0.996	0.999	0.998
Slope	-3.512	-3.235	-3.194	-3.309	-3.246	-3.235
Intercept	40.24	38.46	40.08	39.39	38.30	39.47

Table 6: Intraassay variation shown as mean C_q value $\pm$ standard deviation (SD) from three replicate singleplex assays and three replicate multiplex assays of serial dilution of plasmids from 10^1 to 10^6 copies/μL. Efficiency, square of correlation coefficient (R^2), slope, and intercept of corresponding standard curves are shown.

than or equal to 10 copies/μL. Then, tests were performed to determine LOD, the concentration at which 95% of the positive samples are detected [80]. Six replicates of 10^0 to 10^3 plasmid copies per reaction were used to determine the

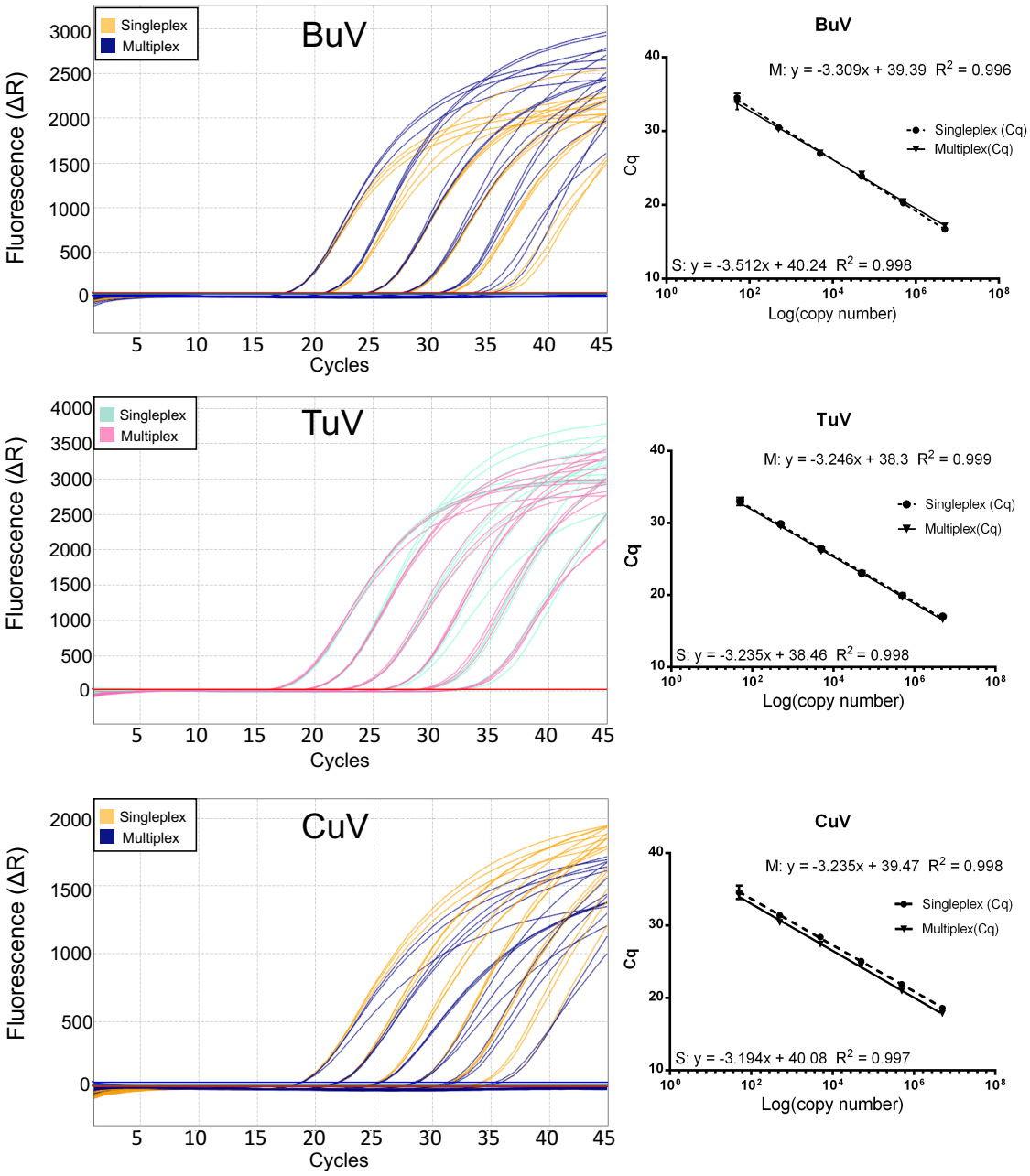

Figure 10: The right panel shows qPCR results of serial dilutions of BuV, TuV, and CuV plasmids (10^1 to 10^6 copies/μL) with baseline-corrected fluorescence (y axis) plotted against cycles (x axis) for singleplex and multiplex assays. The left panel shows corresponding standard curves with the C_q value (y) plotted against the logarithm of the plasmids copies (x) for both singleplex (S; dash line) and multiplex (M; solid line) assays.

LOD for both multiplex and singleplex assays. Probit link function was used to fit a generalized linear model to the observed proportion of positive results and to the logarithmic numbers of plasmid copies by using the MATLAB softwares's(Mathworks, Natick, MA) GLMFIT command (Figure 11) [83, 84]. Profit analysis showed assay sensitivities range from $\sim$ 4 copies/μL (BuV) to $\sim$ 10 copies/μL (TuV) for the singleplex assays, and from $\sim$ 5 copies/μL (TuV) to $\sim$ 9 copies/μL (BuV) for the multiplex assays. In all, the LODs of multiplex assay are not significantly compromised compared to uniplex assays and are $\leqslant$ 10 copies/μL.

5.4 Specificity

In silico BLAST search showed that primers are specific to each target and are unlikely to amplify other undesired sequences. SYBR® Green assay, melting curve analysis indicated that the primers, either in singleplex or multiplex form, were able to amplify specifically the target sequence in regardless of the presence of human DNA. However, when no target sequence but human DNA was present in the reaction, amplification can sometimes be observed at high C_q value (>30 cycles).

The specificity of the hydrolysis probe based qPCR assay were evaluated with cloned full or near full length genomes of parvovirus B19 genotype 1, HBoV1-4, PARV4, and a mix of 13 HPyV mix at 5 x 10^6 copies per reaction. Human DNA extracted from HEK293, HeLa, hFSF, Hep G2, and A549 cells at 90 ng per reaction without template DNA was also tested. None of the specimens was amplified in either the singleplex or multiplex qPCRs assays.

5.5 Cross-reactivity

Since CuV and BuV, especially BuV2, shares high identity in VP region (82% aa identity in VP1) [25], potential cross-amplification of BuV and CuV plasmid template to the two uniplex assays were studied at 5 x 10^7 copies per reaction of BuV 1-3 VP2 clones, and CuV-BR-337 VP2 clones. No cross amplification was observed with any of the assays.

In addition, multiplex assay were able to amplify the corresponding plasmids from a mixture of BuV1 NS1, TuV VP2, and CuV VP2 plasmids with three repli-

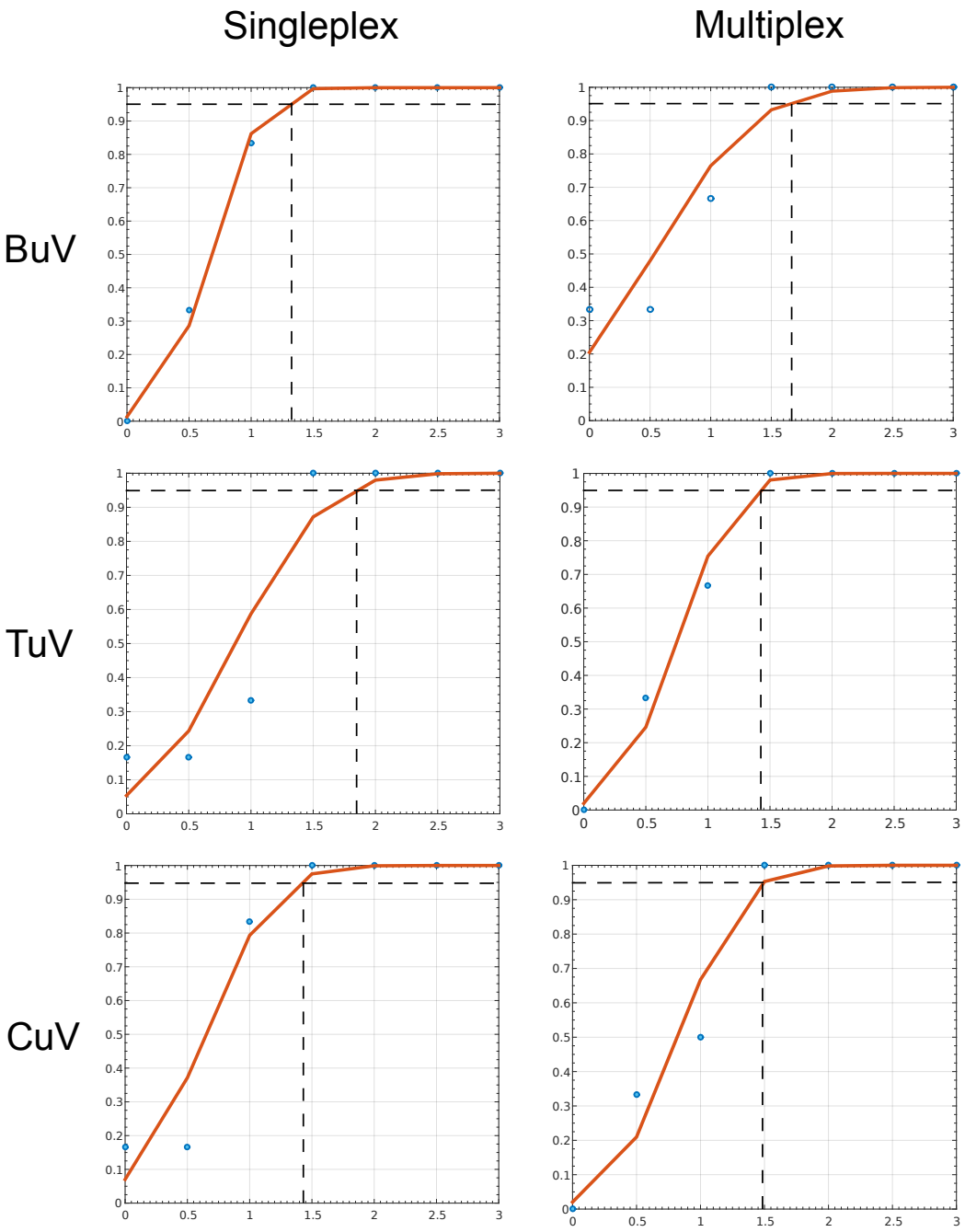

Figure 11: Analytical sensitivity of singleplex and multiplex qPCR assays. X axis denotes the logarithm of the plasmids copies per reaction. A volume of 5 μL templates were added to the reaction. Y axis denotes the proportion of PCR positive samples. Dashed lines indicate the 95% LOD.

cates. Table 7 shows the average C_q value of BuV, TuV, and CuV measured from multiplex qPCR of mix templates and of target templates only reactions.

copies/μL	BuV average C_q		TuV average C_q		CuV average C_q	
	Mix	BuV template	Mix	TuV template	Mix	CuV template
10^1	32.16	33.55	34.52	33.60	32.21	31.95
10^2	30.77	28.20	30.40	29.22	28.89	29.60
10^3	27.24	27.20	26.85	25.97	25.94	25.34
10^4	23.96	22.41	23.17	21.68	22.74	22.96
10^5	20.61	20.07	20.35	19.38	19.44	19.56
10^6	17.21	16.61	16.96	16.37	16.18	16.24
10^7	13.94	13.88	13.63	13.62	13.07	13.19

Table 7: DNA quantification from heterologous mixtures of BuV, TuV and CuV plasmids versus quantification from homologous template.

5.6 Repeatability and reproducibility

Intraassay variation (short-term repeatability) was studied with three replicates of plasmid dilution series from $10^1/\mu$L to $10^6/\mu$L within in a single run, and interassay variation (long-term reproducibility) was studied with three replicates, of which each was performed as triplicates on different days. The results are shown as standard deviations (SD) and coefficients of variations (CV%s) to represent intraassay variation (Table 8 and Table 10) and interassay variation (Table 9 and Table 10).

CV is the ratio of standard deviation to the mean. As recommended by Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) guidelines, the CV should not be calculated from C_q values [80, 85] but can be calculated with copy numbers or concentrations. In this study, CV%s for intra- and interassay variation were calculated from mean copy number of each plasmid dilution, using the formula below:

$$CV\% = \frac{SD}{\bar{x}} = \frac{\sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}}{\bar{x}} \times 100\%$$

where x_i denotes the mean measured quantity of DNA, $\bar{x}$ denotes the overall mean of replicates, and n denotes the number of replicates (in this case, equals to three).

copies/μL	Intraassay variation (Mean C_q value $\pm$ SD)					
	Singleplex			Multiplex		
	BuV	TuV	CuV	BuV	TuV	CuV
10^1	34.55 ± 0.59	33.01 ± 0.57	34.58 ± 0.92	33.84 ± 0.91	32.85 ± 0.21	34.33 ± 0.26
10^2	30.52 ± 0.12	29.86 ± 0.24	31.42 ± 0.09	30.28 ± 0.13	29.61 ± 0.25	30.42 ± 0.27
10^3	26.98 ± 0.05	26.42 ± 0.26	28.41 ± 0.02	27.04 ± 0.13	26.17 ± 0.02	27.42 ± 0.08
10^4	23.88 ± 0.15	23.06 ± 0.12	25.09 ± 0.03	24.18 ± 0.05	22.91 ± 0.14	24.40 ± 0.07
10^5	20.29 ± 0.09	19.91 ± 0.05	21.88 ± 0.12	20.46 ± 0.06	19.76 ± 0.11	20.94 ± 0.08
10^6	16.72 ± 0.12	17.02 ± 0.05	18.60 ± 0.06	17.14 ± 0.02	16.69 ± 0.17	17.84 ± 0.06
Efficiency	92.64%	103.78%	105.63%	100.53%	103.25%	103.76%
R^2	0.998	0.998	0.997	0.996	0.999	0.998
Slope	-3.512	-3.235	-3.194	-3.309	-3.246	-3.235
Intercept	40.24	38.46	40.08	39.39	38.30	39.47

Table 8: Intraassay variation shown as mean C_q value $\pm$ standard deviation (SD) from three replicate singleplex assays and three replicate multiplex assays of serial dilution of plasmids from 10^1 to 10^6 copies/μL. Efficiency, square of correlation coefficient (R^2), slope, and intercept of corresponding standard curves are shown.

copies/μL	Interassay variation (Mean C_q value $\pm$ SD)					
	Singleplex			Multiplex		
	BuV	TuV	CuV	BuV	TuV	CuV
10^1	35.05 ± 0.82	32.81 ± 0.58	34.22 ± 0.42	33.74 ± 0.13	32.77 ± 0.35	33.10 ± 1.10
10^2	31.11 ± 0.59	29.37 ± 0.47	31.48 ± 0.32	30.61 ± 0.35	28.96 ± 0.93	29.32 ± 0.95
10^3	27.44 ± 0.43	26.06 ± 0.44	28.13 ± 0.29	26.77 ± 0.94	25.82 ± 0.30	26.71 ± 0.62
10^4	24.44 ± 0.51	22.67 ± 0.38	25.08 ± 0.53	24.18 ± 0.63	22.67 ± 0.36	23.34 ± 1.00
10^5	20.89 ± 0.85	19.55 ± 0.48	21.75 ± 0.40	20.61 ± 0.65	19.34 ± 0.47	20.39 ± 0.49
10^6	17.75 ± 1.25	16.67 ± 0.37	18.68 ± 0.54	17.05 ± 0.12	15.94 ± 0.67	17.05 ± 0.68

Table 9: Interassay variation shown as mean C_q value $\pm$ standard deviation (SD) from three replicate singleplex assays and three replicate multiplex assays of serial dilution of plasmids from 10^1 to 10^6 copies/μL.

copies/μL	Intraassay variation (CV%)						Interassay variation (CV%)					
	Singleplex			Multiplex			Singleplex			Multiplex		
	BuV	TuV	CuV	BuV	TuV	CuV	BuV	TuV	CuV	BuV	TuV	CuV
10^1	39.83	35.45	51.51	55.15	18.52	14.31	43.10	9.86	17.61	9.55	29.53	13.48
10^2	7.69	16.57	6.47	9.27	19.76	17.04	10.88	13.20	14.33	22.86	28.92	10.76
10^3	2.96	17.45	1.51	9.21	5.88	1.43	15.12	0.54	19.73	16.47	17.22	15.58
10^4	9.80	8.02	2.52	3.18	5.05	9.30	21.55	7.86	11.35	11.51	17.10	17.95
10^5	6.19	3.34	8.74	4.18	5.57	7.73	10.70	8.35	8.09	10.49	9.65	14.70
10^6	7.52	3.56	4.01	1.45	4.30	12.31	15.55	2.23	3.64	6.93	10.37	9.40

Table 10: CV%s for intra- and interassay variations were determined from three replicates performed on a single run and on different days, respectively. CV% was calculated from mean measured DNA copy number of each plasmid dilution rather than C_q values.

The highest intraassay variations of singleplex and multiplex assays were both observed at 10^1 copies/μL, with mean SD of 0.69 and 0.46, respectively. The highest interassay variation of singleplex and multiplex assays were observed at the lowest quantity, with a mean of CV% of 23.52% and 17.52%, respectively.

5.7 qPCR of skin biopsy DNA extracts

5.7.1 Immunosuppressed cohort

To evaluate the qPCR method on clinical specimens, skin biopsies collected from transplant patients in Finland were screened for *Protoparvovirus* DNA. Four of the 137 transplant patients (2.92%) were found CuV DNA positive (Table 11).

Of the 131 liver-transplant patients, three (2.3%) were found reproducibly positive for CuV DNA with quantity ranging from 2711.96 to 4506.04 copies per million cells. DNA extract of two sera from PCR-positive individuals (P-022 and P-065) were also qPCR tested, the results of which were negative. Serum of patient P-108 was not available.

Of the five kidney-transplant patients, one (20%) was found positive for CuV DNA. This patient was diagnosed with squamous cell carcinoma. Three skin biopsies were taken from him: one from healthy skin, one from carcinoma *in situ* (Bowen's disease, premalignant stage of carcinoma) area, and one from carcinoma area; all were positive for CuV DNA. The most abundant amount of CuV DNA (10032.29 copies per million cells) was found in carcinoma tissue and the least (254.48 copies per million cells) in healthy tissue.

All four patients received their transplant more than five years before the samples were taken and were all under immunosuppressive treatment. Further patient characteristics, including the time of transplant and sampling, immunosuppressants and other diseases can be found in Table 11.

Patient ID	Transplant organs	Transplant time	Sampling time	Gender	DNA load (copies per 10^6 cells)	Immuno-suppressant	Other diseases
P-022[a]	Liver	67 (2003)	77 (2013)	Male	4506.043	Advagraf, CellCept	no other diseases
P-065[a]	Liver	47 (1991)	68 (2012)	Male	3550.362	Sandimmun, Medrol	ulcerative colitis
P-108	Liver	52 (2009)	57 (2014)	Female	2711.962	Medrol, CellCept, Advagraf	diabetes, hypothyreoiditis, hypertension arterialis
A	Kidney	50 (1986) 61 (1997)	–	Male	254.4841[b] 2280.92[c] 10032.29[d]	Sandimmun	blood pressure, coronary disease

Table 11: Characteristics of transplant patients with CuV DNA detected in skin tissue.
[a] Serum samples were also qPCR tested, results of which were negative.
[b] The sample was taken from healthy skin area.
[c] The sample was taken from carcinoma *in situ* area.
[d] The sample was taken from carcinoma area.

5.7.2 Non-immunosuppressed cohort

93 non-immunosuppressed adults with allergic and/or irritant contact dermatitis were treated as control for the purpose of this study. 144 skin punch biopsies were taken from allergic or irritant reaction area, none of which was positive for BuV, TuV, or CuV DNA.

5.7.3 BuV and TuV positive samples

BuV and TuV DNA positive supernatants from the original discovery studies were used as positive controls to validate the multiplex qPCR performance on detecting BuV and TuV DNA. The results are shown in Table 12. The qPCR results of BuV DNA quantity was compared with that from a qPCR performed in 2012 with the same singleplex BuV assay [66]. All samples have been stored at $-20°$C,

Virus	copies/mL (2012)	copies/mL (2017)
BuV1-BF7	4.94×10^3	1.72×10^3
BuV1-BF86	8.12×10^3	1.93×10^3
BuV1-BF96	4.95×10^4	1.38×10^4
BuV2-BF39	3.74×10^3	2.90×10^3
TuV	–	4.22×10^4

Table 12: qPCR quantification of BuV and TuV DNA positive samples. The quantification in 2012 was performed with the same BuV singleplex assay. The quantification in 2017 was performed with the multiplex assay.

but several freeze and thaw cycles have taken place on these samples during the past five years.

6 Discussion and prospects

In this study, real-time singleplex assays and a multiplex assay were developed to facilitate the detection and quantification of BuV, TuV, and CuV DNA and to assess the prevalence of them among different groups. The assay was able to detect all three human-associated protoparvoviruses simultaneously in one single step.

The multiplex qPCR offers several advantages over the singleplex qPCR in: 1) efficiency, as less sample and time are required to acquire more information; 2) economy, as less reagents are used for the amplification of multiple targets; 3) accuracy, as fewer pipetting errors are introduced since less pipetting is performed. However, there are some drawbacks regarding the multiplex of uniquely-labeled probes. In the assay, FAM, Texas Red, and JOE dyes were utilized for the detection. Thus, a real-time PCR machine that has multiplexing capability with multiple detection channels for these three dyes is required. Another requirement for the real-time PCR system is the independency of a passive reference dye for fluorescence data normalization, as ROX was excluded in this assay due to its overlapping emission spectrum with that of Texas Red. In all, the multiplex of fluorescent dyes may be a challenge for some laboratories with qPCR machine that does not fulfil the above-mentioned requirements.

An estimated 119,873 solid organ transplants were performed in 2014 worldwide [86] and around 2,000 organ transplants were performed yearly in Nordic countries [87]. To prevent organ rejection, immunosuppressive agents are given to transplant recipients to suppress their adaptive and innate immune responses [88]. As a result, this population is more likely to be infected and affected by "opportunists", such as bacterial, viral, and fungal organisms that rarely infect or affect non-immunosuppressed population. Therefore, we evaluated the prevalence of these viruses among 137 immunosuppressed transplant recipients and 93 non-immunosuppressed individuals. None of BuV, TuV, nor CuV DNA was detected in the non-immunosuppressed cohort. While the DNA of BuV and TuV was also absent from the skin of the immunosuppressed cohort, DNA of CuV was detected in the skin of four immunosuppressed individuals.

This is the first study to investigate the presence and quantity of BuV and TuV in skin tissues. The absence of BuV and TuV in skin tissues indicates the preferable replication sites may not be in skin cells. So far, BuVs has been predominately

found in fecal samples with once in nasal swab, and TuV has only been detected in one fecal sample. More evidence is needed to determine whether TuV is a human virus.

This study showed that the presence CuV is not limited to patients with CTCL or cutaneous malignant melanoma but also immunocompromised individuals with healthy skin or squamous cell carcinoma. Among the four CuV DNA positive individuals, three were liver transplant recipients and one was kidney transplant recipient. The sera of two of the CuV positive liver transplant recipients were further examined with the multiplex qPCR. CuV DNA was not detected in the serum samples, indicating the infection may be local and may not result in a viremia, or the amount of DNA in serum was too low to be detected. A coupled serological study would further our understanding of the pathogenesis of CuV. One limitation of the study is the lack of healthy skin of healthy individuals. However, as the skin biopsy is an invasive sampling procedure, the availability of such samples would be limiting.

Parvoviruses are generally considered oncolytic rather than oncogenic [25, 30]. The results showed that the quantity of CuV DNA was higher in cancerous area but lower in the normal specimen. This difference in quantity may be due to the presence of more rapid dividing cells in S phase in the tumors for viral replication [89]. The detection of CuV DNA in low quantity in both healthy and cancerous skin of immunosuppressed may argue its oncogenic role in skin cancers. Previous study has shown parvovirus B19 persists in healthy skin for lifelong time after primary infection [90]. As is the case with B19, the presence of CuV DNA might be a result of persistent infection. However, many human parvoviruses have been identified in tumors. B19 DNA has been detected in thyroid carcinoma [91, 92] and B19 capsid proteins were detected in lymphoma [93]. The existence of human bocavirus DNA in tumor samples of lung or colorectal cancer patients was reported [94, 95]. AAV-2 was found to induce insertions to known oncegenes in human hepatocellular carcinomas [96]. Nevertheless, more studies are needed to address its role in cancers.

Members of the *Rodent protoparvovirus 1* species in the *Protoparvovirus* genus are known oncolytic parvoviruses, which include *Rodent protoparvoviruses* H-1, the minute virus of mice and LuIII. They have been found to be efficient at infecting and killing tumor cells, while sparing normal cells [22]. H-1 parvovirus has already been manipulated as anticancer agent to treat glioblastoma multiforme

[22]. Tumor virus X, isolated from a permanent human amnion cell line, also exhibited oncolytic potentials [97]. Whether CuV possesses oncolytic properties requires both *in vivo* and *in vitro* investigations.

In conclusion, a multiplex qPCR assay that is high in sensitivity (LOD $\leqslant$ 10 copies/μL), specificity and reproducibility has been developed in this study to detect and quantify three emerging human-associated protoparvoviruses, and to facilitate the future studies of their prevalence and pathobiology. The absence of BuV and TuV DNA in skin tissues indcates skin cells may not be the preferable replication sites. And further data are needed to determine whether TuV is a human virus. The findings further confirmed the existence of CuV DNA in or on human skin and posed questions for its potential role in skin cancers.

References

1 Gerba, C. P, Rose, J. B, & Haas, C. N. (1996) Sensitive population: who is at the greatest risk? *International Journal of Food Microbiology* **30**, 113–123.

2 Wylie, K. M, Weinstock, G. M, & Storch, G. A. (2012) Emerging view of the human virome. *Translational Research* **160**, 283 – 290. In-Depth Review: Of Microbes and Men: Challenges of the Human Microbiome.

3 Minot, S, Bryson, A, Chehoud, C, Wu, G. D, Lewis, J. D, & Bushman, F. D. (2013) Rapid evolution of the human gut virome. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 12450 – 12455.

4 Norman, J. M, Handley, S. A, Baldridge, M. T, Droit, L, Liu, C. Y, Keller, B. C, Kambal, A, Monaco, C. L, Zhao, G, Fleshner, P, Stappenbeck, T. S, McGovern, D. P, Keshavarzian, A, Mutlu, E. A, Sauk, J, Gevers, D, Xavier, R. J, Wang, D, Parkes, M, & Virgin, H. W. (2015) Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell* **160**, 447 – 460.

5 Monaco, C. L, Gootenberg, D. B, Zhao, G, Handley, S. A, Ghebremichael, M. S, Lim, E. S, Lankowski, A, Baldridge, M. T, Wilen, C. B, Flagg, M, Norman, J. M, Keller, B. C, Luévano, J. M, Wang, D, Boum, Y, Martin, J. N, Hunt, P. W, Bangsberg, D. R, Siedner, M. J, Kwon, D. S, & Virgin, H. W. (2016) Altered virome and bacterial microbiome in human immunodeficiency virus-associated acquired immunodeficiency syndrome. *Cell Host & Microbe* **19**, 311 – 322.

6 Breitbart, M & Rohwer, F. (2005) Methods for discovering novel DNA viruses in blood using viral particle selection and shotgun sequencing. *BioTechniques* **39**, 729–736.

7 Furuta, R. A, Sakamoto, H, Kuroishi, A, Yasiui, K, Matsukura, H, & Hirayama, F. (2015) Metagenomic profiling of the viromes of plasma collected from blood donors with elevated serum alanine aminotransferase levels. *Transfusion* **55**, 1889–1899.

8 Moustafa, A, Xie, C, Kirkness, E, Biggs, W, Wong, E, Turpaz, Y, Bloom, K, Delwart, E, Nelson, K. E, Venter, J. C, & Telenti, A. (2017) The blood DNA virome in 8,000 humans. *PLOS Pathogens* **13**, 1–20.

9 Foulongne, V, Sauvage, V, Hebert, C, Dereure, O, Cheval, J, Gouilh, M. A, Pariente, K, Segondy, M, Burguière, A, Manuguerra, J.-C, Caro, V, & Eloit, M. (2012) Human skin microbiota: High diversity of DNA viruses identified on the human skin by high throughput sequencing. *PLOS ONE* **7**, 1–11.

10 Hannigan, G. D, Meisel, J. S, Tyldsley, A. S, Zheng, Q, Hodkinson, B. P, SanMiguel, A. J, Minot, S, Bushman, F. D, & Grice, E. A. (2015) The human skin double-stranded DNA virome: Topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome. *mBio* **6**.

11 Oh, J, Byrd, A. L, Park, M, Kong, H. H, & Segre, J. A. (2016) Temporal stability of the human skin microbiome. *Cell* **165**, 854 – 866.

12 Canuti, M, Deijs, M, Jazaeri Farsani, S. M, Holwerda, M, Jebbink, M. F, de Vries, M, van Vugt, S, Brugman, C, Verheij, T, Lammens, C, Goossens, H, Loens, K, Ieven, M, & van der Hoek, L. (2014) Metagenomic analysis of a sample from a patient with respiratory tract infection reveals the presence of a γ-papillomavirus. *Frontiers in Microbiology* **5**, 347.

13 Wylie, K. M, Mihindukulasuriya, K. A, Zhou, Y, Sodergren, E, Storch, G. A, & Weinstock, G. M. (2014) Metagenomic analysis of double-stranded DNA viruses in healthy adults. *BMC Biology* **12**, 71.

14 Young, J, Chehoud, C, Bittinger, K, Bailey, A, Diamond, J, Cantu, E, Haas, A, Abbas, A, Frye, L, Christie, J, Bushman, F, & Collman, R. (2015) Viral metagenomics reveal blooms of Anelloviruses in the respiratory tract of lung transplant recipients. *American Journal of Transplantation* **15**, 200–209.

15 Vicente, D, Cilla, G, Montes, M, Pérez-Yarza, E. G, & Pérez-Trallero, E. (2007) Human bocavirus, a respiratory and enteric virus. *Emerging infectious diseases* **13**, 636.

16 Edner, N, Castllo-Rodas, P, Falk, L, Hedman, K, Söderlund-Venermo, M, & Allander, T. (2012) Life-threatening respiratory tract disease with human bocavirus-1 infection in a 4 year-old child. *Journal of Clinical Microbiology* **50**, 531–522.

17 Kernbauer, E, Ding, Y, & Cadwell, K. (2014) An enteric virus can replace the beneficial function of commensal bacteria. *Nature* **516**, 94–98.

18 Bhattarai, N & Stapleton, J. T. (2012) GB virus C: the good boy virus? *Trends in Microbiology* **20**, 124 – 130.

19 Schwarze-Zander, C, Blackard, J. T, & Rockstroh, J. K. (2012) Role of GB virus C in modulating HIV disease. *Expert Review of Anti-infective Therapy* **10**, 563–572.

20 Samulski, R. J & Muzyczka, N. (2014) AAV-mediated gene therapy for research and therapeutic purposes. *Annual Review of Virology* **1**, 427 – 451.

21 (2015) Oncolytic immunotherapy through tumor-specific translation and cytotoxicity of poliovirus. *Discovery Medicine* **19**, 359 – 365.

22 Marchini, A, Bonifati, S, Scott, E. M, Angelova, A. L, & Rommelaere, J. (2015) Oncolytic parvoviruses: from basic virology to clinical applications. *Virology Journal* **12**.

23 Phan, T. G, Vo, N. P, Bonkoungou, I. J. O, Kapoor, A, Barro, N, O'Ryan, M, Kapusinszky, B, Wang, C, & Delwart, E. (2012) Acute diarrhea in West African children: Diverse enteric viruses and a novel parvovirus genus. *Journal of Virology* **86**, 11024–11030.

24 Phan, T. G, Sdiri-Loulizi, K, Aouni, M, Ambert-Balay, K, Pothier, P, Deng, X, & Delwart, E. (2014) New parvovirus in child with unexplained diarrhea, Tunisia. *Emerging Infectious Disease* **20**, 1911–1913.

25 Phan, T. G, Dreno, B, Charlys da Costa, A, Li, L, Orlandi, P, Deng, X, Kapusinszky, B, Siqueira, J, Knol, A.-C, Halary, F, Dantal, J, A. Alexander, K, A. Pesavento, P, & Delwart, E. (2016) A new protoparvovirus in human fecal samples and cutaneous T cell lymphomas (mycosis fungoides). *Virology* **496**, 299 – 305.

26 Carter, J & Saunders, V. (2013) *Virology: Principles and Applications, 2nd Edition*. (John Wiley & Sons, Inc.), 2 edition.

27 Qiu, J, Söderlund-Venermo, M, & Young, N. S. (2017) Human parvoviruses. *Clinical Microbiology Reviews* **30**, 43–113.

28 Kilham, L & Olivier, L. (1959) A latent virus of rats isolated in tissue culture. *Virology* **7**, 428 – 437.

29 Brailovsky, C. (1966) Research on the rat K virus (Parvovirus ratti). i. A method of titration by plaques and its application to the study of the multiplication cycle of the virus. *Annales de I'Institut Pasteur (Paris)* **110**, 49–59.

30 Kerr, J. R, Cotmore, S. F, Bloom, M. E, Linden, R. M, & Parrish, C. R. (2006) *Parvoviruses*. (Hodder Arnold Publication).

31 Parrish, C. R. (1990) Emergence, natural history, and variation of canine, mink, and feline parvoviruses. *Advances in Virus Research* **38**, 403 – 450.

32 Shackelton, L. A, Parrish, C. R, Truyen, U, & Holmes, E. C. (2005) High rate of viral evolution associated with the emergence of carnivore parvovirus. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 379 – 384.

33 Parrish, C. R & Kawaoka, Y. (2005) The origins of new pandemic viruses: The acquisition of new host ranges by canine parvovirus and influenza A viruses.

34 Pattison, J, Jones, S, J., H, Davis, L, White, J, Stroud, C, & Murtaza, L. (1981) Parvovirus infections and hypoplastic crisis in sickle-cell anaemia. *The Lancet* **317**, 664 – 665.

35 Cossart, Y, Cant, B, Field, A, & Widdows, D. (1975) Parvovirus-like particles in human sera. *The Lancet* **305**, 72 – 73.

36 Anderson, M, Jones, S, Fisher-Hoch, S, Lewis, E, Hall, S, Bartlett, C, Cohen, B, Mortimer, P, & Pereira, M. (1983) Human parvovirus, the cause of erythema infectiosum (fifth disease)? *The Lancet* **321**, 1378.

37 Anderson, M, Lewis, E, Kidd, I, Hall, S, & Cohen, B. (1984) An outbreak of erythema infectiosum associated with human parvovirus infection. *Journal of Hygiene (London)* **93**, 85–93.

38 Brown, T, Ritchie, L, Clewley, J, & Reid, T. (1984) Intrauterine parvovirus infection associated with hydrops fetalis. *The Lancet* **2**, 1033–1034.

39 Knipe, D. M & Howley, P. M. (2007) *Fields' Virology*. (Wolters Kluwer Health \Lippincott Williams & Wilkins), 5 edition.

40 Cotmore, S. F, Agbandje-McKenna, M, Chiorini, J. A, Mukha, D. V, Pintel, D. J, Qiu, J, Söderlund-Venermo, M, Tattersall, P, Tijssen, P, Gatherer, D, & Davison, A. J. (2014) The family Parvoviridae. *Archives of Virology* **159**, 1239–1247.

41 (2017) Virus Taxonomy: The Online (10th) Report of the ICTV.

42 van Regenmortel, M. H, Fauquet, C. M, Bishop, D. H, Carsten, E, Estes, M, Lemon, S, Maniloff, J, Mayo, M, McGeoch, D, Pringle, C, & Wickner, R. (2000) *Virus Taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses*. (Academic Press), 1 edition.

43 Thompson, J. D, Higgins, D. G, & Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673 – 4680.

44 Larkin, M, Blackshields, G, Brown, N, Chenna, R, McGettigan, P, McWilliam, H, Valentin, F, Wallace, I, Wilm, A, Lopez, R, Thompson, J, Gibson, T, & Higgins, D. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947.

45 Dayhoff, M. O, Schwartz, R. M, & Orcutt, B. (1978) *Atlas of Protein Sequence and Structure*.

46 Saitou, N & Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular biology and evolution* **4**, 406–425.

47 Kumar, S, Stecher, G, & Tamura, K. (2016) Mega7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870 –1874.

48 Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783 – 791.

49 Nei, M & Kumar, S. (2000) *Molecular Evolution and Phylogenetics*. (Oxford University Press), 1 edition.

50 Rascovan, N, Duraisamy, R, & Desnues, C. (2016) Metagenomics and the human virome in asymptomatic individuals. *Annual Review of Microbiology* **70**, 125–141.

51 Delwart, E. L. (2007) Viral metagenomics. *Reviews in Medical Virology* **17**, 115–131.

52 Fontana, J. M, Alexander, E, & Salvatore, M. (2012) Translational research in infectious disease: current paradigms and challenges ahead. *Translational Research* **159**, 430 – 453.

53 Williams, S. C. (2013) The other microbiome. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 2682–2684.

54 Jones, M. S, Kappor, A, Lukashov, V. V, Simmonds, P, Hecht, F, & Delwart, E. (2005) New DNA viruses identified in patients with acute viral infection syndrome.

55 Kappor, A, Slikas, E, Simmonds, P, Chieochansin, T, Naeem, A, Shaukat, S, Alam, M. M, Sharif, S, Angez, M, Zaidi, S, & Delwart, E. (2009) A newly identified bocavirus species in human stool. *The Journal of Infectious Diseases* **199**, 196–200.

56 Kappor, A, Simmonds, P, Slikas, E, Li, L, Bodhidatta, L, Senthabutr, O, Triki, H, Bahri, O, Oderinde, B. S, Baba, M. M, Bukbuk, D. N, Besser, J, Bartkus, J, & Delwart, E. (2010) Human bocaviruses are highly diverse, dispersed, recombination prone, and prevalent in enteric infections. *The Journal of Infectious Diseases* **201**, 1633–1643.

57 Kantola, K. (2014) Ph.D. thesis (University of Helsinki).

58 Yahiro, T, Wangchuk, S, Tshering, K, Bandhari, P, Zangmo, S, Dorji, T, Tshering, K, Matsumoto, T, Nishizono, A, Söderlund-Venermo, M, & Ahmed, K. (2014) Novel human bufavirus genotype 3 in children with severe diarrhea, Bhutan. *Emerging Infectious Disease* **20**, 1037 – 1039.

59 Tattersall, P & Ward, D. C. (1976) Rolling hairpin model for replication of parvovirus and linear chromosomal DNA. *Nature* **263**, 106–109.

60 Hulo, C, de Castro, E, Masson, P, Bougueleret, L, Bairoch, A, Xenarios, I, & Le Mercier, P. (2011) ViralZone: a knowledge resource to understand virus diversity. *Nucleic Acids Research* **39**, D576.

61 Ayouni, S, Estienney, M, Hammami, S, Neji Guediche, M, Pothier, P, Aouni, M, Belliot, G, & de Rougemont, A. (2016) Cosavirus, salivirus and bufavirus in diarrheal Tunisian infants. *PLOS ONE* **11**, 1–11.

62 Chieochansin, T, Vutithanachot, V, Theamboonlers, A, & Poovorawan, Y. (2015) Bufavirus in fecal specimens of patients with and without diarrhea in Thailand. *Archives of Virology* **160**, 1781–1784.

63 Altay, A, Yahiro, T, Bozdayi, G, Matsumoto, T, Sahin, F, Ozkan, S, Nishizono, A, Söderlund-Venermo, M, & Ahmed, K. (2015) Bufavirus genotype 3 in Turkish children with severe diarrhoea. *Clinical Microbiology and Infection* **21**, 965.e1 – 965.e4.

64 Huang, D.-D, Wang, W, Lu, Q.-B, Zhao, J, Guo, C.-T, Wang, H.-Y, Zhang, X.-A, Tong, Y.-G, Liu, W, & Cao, W.-C. (2015) Identification of bufavirus-1 and bufavirus-3 in feces of patients with acute diarrhea, China. *Scientific Reports* **5**, 13272 EP –.

65 Väisänen, E, Paloniemi, M, Kuisma, I, Lithovius, V, Kumar, A, Franssila, R, Ahmed, K, Delwart, E, Vesikari, T, Hedman, K, & Söderlund-Venermo, M. (2016) Epidemiology of two human protoparvoviruses, bufavirus and tusavirus. *Scientific Reports* **6**, 39267 EP –.

66 Väisänen, E, Kuisma, I, Phan, T. G, Delwart, E, Lappalainen, M, Tarkka, E, Hedman, K, & Söderlund-Venermo, M. (2014) Bufavirus

in Feces of patients with gastroenteritis, Finland. *Emerging Infectious Diseases* **20**, 1077–1079.

67 Smits, S. L, Schapendonk, C. M, van Beek, J, Vennema, H, C. Schürch, A. C, Schipper, D, Bodewes, R, Haagmans, B. L, Osterhaus, A. D. M. E, & Koopmans, M. P. (2014) New viruses in idiopathic human diarrhea cases, the Netherlands. *Emerging Infectious Diseases* **20**, 1218 – 1222.

68 Handley, S. A, Thackray, L. B, Zhao, G, Presti, R, Miller, A. D, Droit, L, Abbink, P, Maxfield, L. F, Kambal, A, Duan, E, Stanley, K, Kramer, J, Macri, S. C, Permar, S. R, Schmitz, J. E, Mansfield, K, Brenchley, J. M, Veazey, R. S, Stappenbeck, T. S, Wang, D, Barouch, D. H, & Virgin, H. W. (2012) Pathogenic simian immunodeficiency virus infection is associated with expansion of the enteric virome. *Cell* **151**, 253–266.

69 Sasaki, M, Orba, Y, Anindita, P. D, Ishii, A, Ueno, K, Hang'ombe, B. M, Mweene, A. S, Ito, K, & Sawa, H. (2015) Distinct lineages of bufavirus in wild shrews and nonhuman primates. *Emerging Infectious Diseases* **21**, 1230–1233.

70 Kemenesi, G, Dallos, B, Görföl, T, Estók, P, Boldogh, S, Kurucz, K, Oldal, M, Marton, S, Bányai, K, & Jakab, F. (2015) Genetic diversity and recombination within bufaviruses: Detection of a novel strain in Hungarian bats. *Infection, Genetics and Evolution* **33**, 288 – 292.

71 Wu, Z, Yang, L, Ren, X, He, G, Zhang, J, Yang, J, Qian, Z, Dong, J, Sun, L, Zhu, Y, Du, J, Yang, F, Zhang, S, & Jin, Q. (2015) Deciphering the bat virome catalog to better understand the ecological diversity of bat viruses and the bat origin of emerging infectious diseases. *The ISME Journal* **10**, 609 – 620.

72 Yang, S, Liu, D, Wang, Y, Qu, F, He, Y, Sun, Z, Shen, Q, Li, W, Fu, X, Deng, X, Zhang, W, & Delwart, E. (2016) Bufavirus protoparvovirus in feces of wild rats in china. *Virus Genes* **52**, 130–133.

73 Hargitai, R, Pankovics, P, Kertész, A. M, Bíró, H, Boros, Á, Phan, T. G, Delwart, E, & Reuter, G. (2016) Detection and genetic characteriza-

tion of a novel parvovirus distantly related to human bufavirus in domestic pigs. *Archives of Virology* **161**, 1033–1037.

74 Liu, L, Schwarz, L, Ullman, K, Ahola, H, Qiu, Y, Ma, Z, & Hennig-Pauka, I. (2016) Identification of a novel bufavirus in domestic pigs by a viral metagenomic approach. *Journal of General Virology* **97**, 1592–1596.

75 Sasaki, M, Gonzalez, G, Wada, Y, Setiyono, A, Handharyani, E, Rahmadani, I, Taha, S, Adiani, S, Latief, M, Kholilullah, Z. A, Subangkit, M, Kobayashi, S, Nakamura, I, Kimura, T, Orba, Y, Ito, K, & Sawa, H. (2016) Divergent bufavirus harboured in megabats represents a new lineage of parvoviruses. *Scientific Reports* **6**, 24257 EP –.

76 Kluge, M, Campos, F. S, Tavares, M, de Amorim, D. B, Valdez, F. P, Giongo, A, Roehe, P. M, & Franco, A. C. (2016) Metagenomic survey of viral diversity obtained from feces of Subantarctic and South American fur seals. *PLOS ONE* **11**, 1–24.

77 Mollerup, S, Fridholm, H, Vinner, L, Kjartansdóttir, K. R, Friis-Nielsen, J, Asplund, M, Herrera, J. A, Steiniche, T, Mourier, T, Brunak, S, Willerslev, E, Izarzugaza, J. M, Hansen, A. J, & Nielsen, L. P. (2017) Cutavirus in cutaneous malignant melanoma. *Emerging Infectious Diseases* **23**, 363–365.

78 Sievers, F, Wilm, A, Dineen, D, Gibson, T. J, Karplus, K, Li, W, Lopez, R, McWilliam, H, Remmert, M, Söding, J, Thompson, J. D, & Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**.

79 Coordinators, N. R. (2013) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research* **41**, D8 –D20.

80 Bustin, S. A, Benes, V, Garson, J. A, Hellemans, J, Huggett, J, Kubista, M, Mueller, R, Nolan, T, Pfaffl, M. W, Shipley, G. L, Vandesompele, J, & Wittwer, C. T. (2009) The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry* **55**, 611–622.

81 Sadeghi, M, Aaltonen, L.-M, Hedman, L, Chen, T, Söderlund-Venermo, & Hedman, K. (2014) Detection of TS polyomavirus DNA in tonsillar tissues of children and adults: Evidence for site of viral latency. *Journal of Clinical Virology* **59**, 55 – 58.

82 Toppinen, M, Norja, P, Aaltonen, L.-M, Wessberg, S, Hedman, L, Söderlund-Venermo, M, & Hedman, K. (2015) A new quantitative PCR for human parvovirus B19 genotypes. *Journal of Virological Methods* **218**, 40 – 45.

83 Burns, M & Valdivia, H. (2008) Modelling the limit of detection in real-time quantitative PCR. *European Food Research and Technology* **226**, 1513–1524.

84 Kantola, K, Sadeghi, M, Antikainen, J, Kirveskari, J, Elwart, E, Hedman, K, & Söderlund-Venermo, M. (2010) Real-time quantitative PCR detection of four human bocaviruses. *Journal of Clinical Microbiology* **48**, 4044–4050.

85 Schmittgen, T. D & Livak, K. J. (2008) Analyzing real-time PCR data by the comparative CT method. *Nature Protocols* **3**, 1101–1108.

86 (2016) WHO-ONT. global observatory on donation and transplantation.

87 (2017) Scandiatransplant.

88 Gangappa, S, Kokko, K. E, Carlson, L. M, Gourley, T, Newell, K. A, Pearson, T. C, Ahmed, R, & Larsen, C. P. (2008) Immune responsiveness and protective immunity after transplantation. *Transplant International* **21**, 293–303.

89 Wolter, S, Richards, R, & Armentrout, R. W. (1980) Cell cycle-dependent replication of the DNA of minute virus of mice, a parvovirus. *Biochimica et Biophysica Acta* **607**, 420–431.

90 Norja, P, Hokynar, K, Aaltonen, L.-M, Chen, R, Ranki, A, Partio, E. K, Kiviluoto, O, Davidkin, I, Leivo, T, Eis-Hübinger, A. M, Schneider, B, Fischer, H.-P, Tolba, R, Vapalahti, O, Vaheri, A, Söderlund-Venermo, M, & Hedman, K. (2006) Bioportofolio: lifelong persistence of variant

and prototypic erythrovirus DNA genome in human tissue. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 7450 – 7453.

91 Wang, J, Zhang, H, H.X., L, Wang, D, Li, Y, Wang, W. Q, He, F, Wang, Z, Yan, Q. G, W., C. L, & Huang, G. (2008) Detection of human parvovirus B19 in papillary thyroid carcinoma. *British Journal of Cancer* **98**, 611–618.

92 Adamson, L. A, Fowler, L. J, Clare-Salzler, M. J, & Hobbs, J. A. (2012) Parvovirus B19 infection in Hashimoto's thyroiditis, papillary thyroid carcinoma, and anaplastic thyroid carcinoma. *Thyroid* **21**, 411–417.

93 Polcz, Monica E.and Adamson, L. A, Lu, X, Chang, M. N, Fowler, L. J, & Hobbs, J. A. (2013) Increased IL-6 detection in adult and pediatric lymphoid tissue harboring parvovirus B19. *Journal of Clinical Virology* **57**, 233 – 238.

94 Schildgen, V, Malecki, M, Tillmann, R.-L, Brockmann, M, & Schildgen, O. (2013) The human bocavirus is associated with some lung and colorectal cancers and persists in solid tumors. *PLOS ONE* **8**.

95 Abdel-Moneim, A. S, El-Fol, H. A, Kamel, M. M, Soliman, A. S. A, Mahdi, E. A, El-Gammal, A. S, & Mahran, T. Z. M. (2016) Screening of human bocavirus in surgically excised cancer specimens. *Archives of Virology* **161**, 2095–2102.

96 Nault, J.-C, Datta, S, Imbeaud, S, Franconi, A, Mallet, M, Couchy, G, Letouze, E, Pilati, C, Verret, B, Blanc, J.-F, Balabaud, C, Calderaro, J, Laurent, A, Letexier, M, Bioulac-Sage, P, Calvo, F, & Zucman-Rossi, J. (2015) Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. *Nature Genetics* **47**, 1187–1193.

97 Vollmers, E. M, D'Abramo, A. J, Cotmore, S. F, & Tattersall, P. (2014) Genome sequence of Tumor Virus X, a member of the genus Protoparvovirus in the Family Parvoviridae. *Genome Announcements* **2**, e00758–14.